

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 13:22:14 ; Search time 1898.09 Seconds
(without alignments)
13097.759 Million cell updates/sec

Title: US-09-898-216-2

Perfect score: 1188
Sequence: 1 GGCTTCGGGAGCGACCGCT.....GGAAGCAGATTTCCTGATT 1188

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pal:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_on:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	1183	99.6	1188	6	AR012692	AR012692 Sequence
2	1160	97.6	1303	9	AF190167	AF190167 Homo sapi
3	1135.2	95.6	1278	9	BC002442	BC002442 Homo sapi
4	1132	95.3	1260	9	AF282596	AF282596 Homo sapi
5	1130.4	95.2	1264	9	AK027405	AK027405 Homo sapi
6	1130.4	95.2	1281	9	BC010152	BC010152 Homo sapi
7	1130	95.1	1262	9	BC014990	BC014990 Homo sapi
8	1120	94.3	1257	9	BC003025	BC003025 Homo sapi
9	867.4	73.0	1253	10	AF232178	AF232178 Mus muscu
10	851.6	71.7	1565	10	BC003425	BC003425 Mus muscu
11	516.6	43.5	691	6	AX261777	AX261777 Sequence
12	515	43.4	518	6	AX193464	AX193464 Sequence
13	455.2	38.3	457	6	AX339946	AX339946 Sequence
14	368.2	31.0	740	6	AX260945	AX260945 Sequence
15	315.6	26.6	186656	2	AC095312	AC095312 Rattus no
16	299.4	25.2	301	6	AX302706	AX302706 Sequence
17	280.8	23.6	1677	8	AF236372	AF236372 Zea mays
18	269.4	22.7	1267	8	AY059109	AY059109 Arabidops
19	269.4	22.7	1503	8	AY034924	AY034924 Arabidops
20	246.6	20.8	12805	1	AE008609	AE008609 Rickettsi
21	240.2	20.2	201050	1	AL646064	AL646064 Ralstonia
22	227.4	19.1	312430	1	RPX002	AJ235221 Rickettsi
23	195	16.4	86765	9	AC004472	AC004472 Homo sapi
24	193.4	16.3	195102	9	AL353795	AL353795 Human DNA
25	193.4	16.3	198829	2	AL391668	AL391668 Homo sapi
26	185.4	15.6	35064	4	AF163772	AF163772 Leishmani
27	185.4	15.6	37852	8	SPBC1665	AL023554 S. pombe c
28	180.6	15.2	253	6	A74641	A74641 Sequence 32
29	180.6	15.2	253	6	A77620	A77620 Sequence 32
30	159.2	13.4	207	6	AX261093	AX261093 Sequence
31	158.2	13.3	44588	2	AC020529	AC020529 Drosophila
32	158.2	13.3	148432	3	AC004642	AC004642 Drosophila
33	158.2	13.3	154840	3	AC099018	AC099018 Drosophila
34	158.2	13.3	300542	3	AE003462	AE003462 Drosophila
35	156.8	13.2	948	1	AF226512	AF226512 Neisseria
36	156.8	13.2	948	1	AF226541	AF226541 Neisseria
37	156.8	13.2	948	6	AX043933	AX043933 Sequence
38	156.8	13.2	948	6	AX043939	AX043939 Sequence
39	156.8	13.2	331801	1	NMA442491	AL162755 Neisseria
40	155.2	13.1	948	1	AF226511	AF226511 Neisseria
41	155.2	13.1	948	1	AF226515	AF226515 Neisseria
42	155.2	13.1	948	1	AF226524	AF226524 Neisseria
43	155.2	13.1	948	1	AF235154	AF235154 Neisseria
44	155.2	13.1	948	1	AF235155	AF235155 Neisseria
45	155.2	13.1	948	1	AF235156	AF235156 Neisseria

ALIGNMENTS

RESULT 1	AR012692	1188 bp	DNA	linear	PAT 05-DEC-1998
LOCUS	AR012692	Sequence 2 from patent US 5763589.			
DEFINITION	AR012692				
ACCESSION	AR012692.1	GI:3971010			
VERSION					
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 1188)				
AUTHORS	Hillman,J.L. and Goll,S.K.				
TITLE	Human membrane protein				
JOURNAL	Patent: US 5763589-A 2 09-JUN-1998;				
FEATURES	Location/Qualifiers				
source	1..1188				
BASE COUNT	268 a 316 c 362 g 237 t				5 others
ORIGIN					

Query Match 99.6%; Score 1183; DB 6; Length 1188;
Best local Similarity 100.0%; Pred. No. 8.9e-271;

Matches 1188: Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Oy	1	GGCTTCTGGAGCNAACCCCTCCGCTGCTCTGTTGTTCCGAGAGTCGCTGCGCGGTGG	60		
Db	1	GGCTTCTGGAGCNAACCCCTCCGCTGCTCTGTTGTTCCGAGAGTCGCTGCGCGGTGG	60		
Oy	61	GAATGCTGGCG	120		
Db	61	GAATGCTGGCG	120		
Oy	121	GGCTTCTGGCG	180		
Db	121	GGCTTCTGGCG	180		
Oy	181	TTCTGTCGCGCAGCAGAGAGCGCTGGGTGTGTGAGCGAATGGCGGATTCACCGGATCTG	240		
Db	181	TTCTGTCGCGCAGCAGAGAGCGCTGGGTGTGTGAGCGAATGGCGGATTCACCGGATCTG	240		
Oy	241	GAGCTGTGTTGACATCTCTCATCCCTGTGTAGACCGGATCCGATATGTGACAGTCTC	300		
Db	241	GAGCTGTGTTGACATCTCTCATCCCTGTGTAGACCGGATCCGATATGTGACAGTCTC	300		
Oy	301	AAGGAAATTTGATCAGCTGCTGAGCAGTCGGCTGTGACTCTGACAAATGTAATCTTG	360		
Db	301	AAGGAAATTTGATCAGCTGCTGAGCAGTCGGCTGTGACTCTGACAAATGTAATCTTG	360		
Oy	361	CAATTCATGATGAGTCTTCTTACCTGCGCATCATGACCTTACAGAGCAAGTACGGTGTG	420		
Db	361	CAATTCATGATGAGTCTTCTTACCTGCGCATCATGACCTTACAGAGCAAGTACGGTGTG	420		
Oy	421	GAGGACCTGAGATATGCTCTCACCCAGCTAGCTCAAAACACCATGATCAAGTCTCGGC	480		
Db	421	GAGGACCTGAGATATGCTCTCACCCAGCTAGCTCAAAACACCATGATCAAGTCTCGGC	480		
Oy	481	AAACTCTCTNTGAGCAAAAGTCTTCGCGGAAACGGAGATCCCTGAATGCCAGATTGGAT	540		
Db	481	AAACTCTCTNTGAGCAAAAGTCTTCGCGGAAACGGAGATCCCTGAATGCCAGATTGGAT	540		
Oy	541	GCCATCAACCAAGCTGCTGACTGCTGGGGTATCCGCTGCTTCGCTTATGATGATCAAGAT	600		
Db	541	GCCATCAACCAAGCTGCTGACTGCTGGGGTATCCGCTGCTTCGCTTATGATGATCAAGAT	600		
Oy	601	ATCCTATGCTCCACCCCGGGTGAAGAGTATATGACAGATGAGTGGAGCAGCGCGG	660		
Db	601	ATCCTATGCTCCACCCCGGGTGAAGAGTATATGACAGATGAGTGGAGCAGCGCGG	660		
Oy	661	AAACGGGACCAAGTCTTCTAGTCTGAGGGAGCCGAGATCGGCCATCAATGTCGAGAA	720		
Db	661	AAACGGGACCAAGTCTTCTAGTCTGAGGGAGCCGAGATCGGCCATCAATGTCGAGAA	720		
Oy	721	GGGAAGAAACAGGCGCCAGATCTGCTCCGAAGCAAGAAAGGCTGAGACGATAAATCAG	780		
Db	721	GGGAAGAAACAGGCGCCAGATCTGCTCCGAAGCAAGAAAGGCTGAGACGATAAATCAG	780		
Oy	781	GCAGAGAGAGAGCCAGTCTGCTGGGAAAGCCCAAGGCTAAAGCTGAAGCTATTCGA	840		
Db	781	GCAGAGAGAGAGCCAGTCTGCTGGGAAAGCCCAAGGCTAAAGCTGAAGCTATTCGA	840		
Oy	841	ATCTGCTGCTGAGCTCTGACACAAATATGAGATGACAGAGCTTCACTGAGTGGCC	900		
Db	841	ATCTGCTGCTGAGCTCTGACACAAATATGAGATGACAGAGCTTCACTGAGTGGCC	900		
Oy	901	GAGCAGTATGACGCGGCTTCTCCAAAGTGGCCAGAGACTCCAAACATATCTACTGCC	960		
Db	901	GAGCAGTATGACGCGGCTTCTCCAAAGTGGCCAGAGACTCCAAACATATCTACTGCC	960		
Oy	961	TCCAAACCTTGGGATGTACACAGCATGTGGCTCAGGCGCATGGGTATATGAGAGCCCTC	1020		
Db	961	TCCAAACCTTGGGATGTACACAGCATGTGGCTCAGGCGCATGGGTATATGAGAGCCCTC	1020		
Oy	1021	ACCAAGAGCCAGTGGCAGAGACTCCAGATCACTCTCCAGTGGAGAGCAGAGATGTC	1080		
Db	1021	ACCAAGAGCCAGTGGCAGAGACTCCAGATCACTCTCCAGTGGAGAGCAGAGATGTC	1080		

Oy	1081	CAGGTACAGATGCAAGTNTGATGAGGAATTTGATCGAGTCAAGATGAGTTAGTGAGC	1140		
Db	1081	CAGGTACAGATGCAAGTNTGATGAGGAATTTGATCGAGTCAAGATGAGTTAGTGAGC	1140		
Oy	1141	TGGGCTTCCGAGAGTCTGGGCAAGAGAGCAGATTTCTGATTT	1188		
Db	1141	TGGGCTTCCGAGAGTCTGGGCAAGAGAGCAGATTTCTGATTT	1188		
RESULT 2					
AF190167					
LOCUS					
DEFINITION					
Homo sapiens membrane associated protein SLP-2 (HUSLP2) mRNA,					
complete cds.					
ACCESSION					
AF190167					
VERSION					
AF190167.1 GI:6456117					
KEYWORDS					
SOURCE					
ORGANISM					
human.					
Homo sapiens					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
REFERENCE					
AUTHORS					
Wang, Y. and Morrow, J.S.					
TITLE					
Identification and characterization of human SLP-2, a novel					
homologue of stomatin (band 7.2b) present in erythrocytes and other					
tissues					
JOURNAL					
MEDLINE					
20179914					
REFERENCE					
AUTHORS					
Wang, Y. and Morrow, J.S.					
TITLE					
Direct Submission					
JOURNAL					
Submitted (25-SEP-1999) Pathology, Yale Medical School, 310 Cedar					
Street, New Haven, CT 06510, USA					
FEATURES					
source					
1..1303					
/organism="Homo sapiens"					
/db_xref="taxon:9606"					
/chromosome="9"					
/tissue_type="heart muscle"					
1..1303					
/gene="HUSLP2"					
64..1134					
/gene="HUSLP2"					
/note="stomatin-like protein 2; widely distributed					
peripheral membrane protein; similar to human erythrocyte					
stomatin and MEC 1 of Caenorhabditis elegans; thought to					
be involved in mechanoreception or lipid domain					
organization"					
/codon_start=1					
/product="membrane associated protein SLP-2"					
/protein_id="AA09142.1"					
/db_xref="GI:6456118"					
/translation="MLARARGTGALLRGLSLASGRAPRRASGLPRNTVVLFPVQ					
EAWVERRMGRFRHLEPGILNLLIPLDIRYVSLKIVINPEGSATVLDNLTQID					
GVLYLRINDPYKASYGVDPKAYVQLQNTMRSLKISLDKVRRESINASIVDA					
INQADQWIRLRLREIKDHYPPRVKSMQVDAERKRRTVLESCTRSAINVA					
EGKROAQLIASAEKAEQINQAAAGSAVLAARAEIRLLAALYQHNQDAASL					
TVAEQVSAFSLKLSKDSNTILLPSPNGDVTSMVAQMGVYALTKRAPVGPDPDSLSG					
SSRDVQGTDAISDELDRVKMS"					
217					
misc_feature					
/gene="HUSLP2"					
/note="alternative translation initiation site"					
misc_feature					
/gene="HUSLP2"					
/note="alternative translation initiation site"					
BASE COUNT					
ORIGIN					
315 a 338 c 378 g 272 t					
Query Match					
Best local Similarity 99.4%; Pred. No. 2, 6e-265;					
Matches 1182; Conservative 0; Mismatches 5; Indels 2; Gaps 2;					

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 15:23:18 ; Search time 1896.25 seconds
(without alignments)
13110.468 Million cell updates/sec

Title: US-09-898-216-2

Perfect score: 1188

Sequence: 1 GGCTTCTGGGACNACCGCT.....GGAAGCAGATTTTCCTGATT 1188

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vl.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
-----	-----	-----	-----	-----	-----

1	1188	100.0	1188	6	AR012692		
2	857	72.1	1257	9	BC003025		
3	857	72.1	1260	9	AF282596		
4	857	72.1	1262	9	BC014990		
5	857	72.1	1278	9	BC002442		
6	857	72.1	1303	9	AF190167		
7	806	67.8	1244	9	AK027405		
8	806	67.8	1281	9	BC010152		
9	454	38.2	457	6	AX339946		
10	365	30.7	518	6	AX193464		
11	331	27.9	691	6	AX261777		
c	12	250	21.0	301	6	AX302706	
13	169	14.2	740	6	AX260945		
c	14	149	12.5	55406	2	AC026501	
c	15	149	12.5	86765	9	AC004472	
c	16	149	12.5	195102	9	AL353795	
17	149	12.5	198829	2	AL391668		
18	84	7.1	207	6	AX261093		
19	74	6.2	253	6	A74641		
20	74	6.2	253	6	A77620		
21	53	4.5	1253	10	AF323178		
22	53	4.5	1565	10	BC003425		
23	53	4.5	124230	10	AC005259		
24	33	2.8	193972	2	AC094506		
c	25	23	1.9	193972	2	AC094506	
26	21	1.8	28958	6	AR044578		
27	21	1.8	28958	6	147768		
28	21	1.8	28958	6	150958		
29	21	1.8	28958	6	163356		
30	21	1.8	28958	6	170387		
31	21	1.8	28958	6	185639		
32	21	1.8	28958	6	188045		
33	21	1.8	28958	6	190322		
34	21	1.8	49377	6	188042		
35	21	1.8	67523	1	SCU24241		
c	36	21	1.8	135005	9	HS860F19	
37	21	1.8	151915	2	AC095033		
c	38	21	1.8	160852	9	AC020915	
c	39	21	1.8	163903	9	AC007387	
40	21	1.8	171964	2	AL451003		
c	41	21	1.8	179307	2	AC104027	
42	21	1.8	185232	2	AC104005		
c	43	21	1.8	186656	2	AC095312	
c	44	21	1.8	214042	2	AC023149	
c	45	21	1.8	221880	2	AC105706	

ALIGNMENTS

RESULT 1

AR012692

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT

ORIGIN

linear PAT 05-DEC-1998

DNA

1188 bp

Sequence 2 from patent US 5763589.

AR012692

AR012692

GI:3971010

Unknown.

Unknown.

Unclassified.

1 (bases 1 to 1188)

Hillman, J.L. and Goli, S.K.

Human membrane protein

Patent: US 5763589-A 2 09-JUN-1998;

Location/Qualifiers

1..1188

/organism="unknown"

268 a

316 c

362 g

237 t

5 others

Query Match 100.0%; Score 1188; DB 6; Length 1188;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1188: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Qy	1	GGCTTCTGGGAGCACCCTCGCTCGCTCGCTGGTTCCGGAGTCCCTCGCGCGTGG	60
Db	1	GGCTTCTGGGAGCACCCTCGCTCGCTCGCTGGTTCCGGAGTCCCTCGCGCGTGG	60
Qy	61	GAATGCTGGCG	120
Db	61	GAATGCTGGCG	120
Qy	121	GGCTTCTGGCG	180
Db	121	GGCTTCTGGCG	180
Qy	181	TTCTGTCGCGCAGCAGGAGCGCTGGTGGTGGAGCGAATGGGCGCATTCACCGGATCTCG	240
Db	181	TTCTGTCGCGCAGCAGGAGCGCTGGTGGTGGAGCGAATGGGCGCATTCACCGGATCTCG	240
Qy	241	GAGCCTGTTTGAACATCCTCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTC	300
Db	241	GAGCCTGTTTGAACATCCTCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTC	300
Qy	301	AAGGAATTTGTCATCAACGTCGCTGAGCAGTCGGCTGTGACTCTCGACAATGTAACTCTG	360
Db	301	AAGGAATTTGTCATCAACGTCGCTGAGCAGTCGGCTGTGACTCTCGACAATGTAACTCTG	360
Qy	361	CAATTCGATGGAGTCTTTACCTCGCATCATGACATGACGACCTTACAAGGCAAGCTACGGTGTG	420
Db	361	CAATTCGATGGAGTCTTTACCTCGCATCATGACATGACGACCTTACAAGGCAAGCTACGGTGTG	420
Qy	421	GAGGACCTGAGTATGCGCTGACCGAGTACGCTCAACAAACCATGAGATCAGAGTCGCG	480
Db	421	GAGGACCTGAGTATGCGCTGACCGAGTACGCTCAACAAACCATGAGATCAGAGTCGCG	480
Qy	481	AACTCTCTNTGGCAAAAGTCTTCGGGAACGGAGTCCCTGAAATGCCAGCATTCGGAT	540
Db	481	AACTCTCTNTGGCAAAAGTCTTCGGGAACGGAGTCCCTGAAATGCCAGCATTCGGAT	540
Qy	541	CCATCAACCAAGCTGTGACTGTGGGTATCGCGTGCCTTCGCTTATGAGATCAAGAT	600
Db	541	CCATCAACCAAGCTGTGACTGTGGGTATCGCGTGCCTTCGCTTATGAGATCAAGAT	600
Qy	601	ATCATGTGCCACCGCGGTGAAGAGTCTATGAGATGAGTGGAGGAGAGCGGCGG	660
Db	601	ATCATGTGCCACCGCGGTGAAGAGTCTATGAGATGAGTGGAGGAGAGCGGCGG	660
Qy	661	AAACGGGACAGTCTAGAGTCTGAGGAGCCGAGAGTCCGCCATCAATGTGGCAGAA	720
Db	661	AAACGGGACAGTCTAGAGTCTGAGGAGCCGAGAGTCCGCCATCAATGTGGCAGAA	720
Qy	721	GGGAAGAAACAGGCGGAGTCTTCCGCTCCGAGCAGAAAAGGCTGAACAGATAATCAG	780
Db	721	GGGAAGAAACAGGCGGAGTCTTCCGCTCCGAGCAGAAAAGGCTGAACAGATAATCAG	780
Qy	781	GCAGCAGAGAGGCGGAGTCTGCGGAGGAGCCGAGGCTAAGCTGAAGCTATTGGA	840
Db	781	GCAGCAGAGAGGCGGAGTCTGCGGAGGAGCCGAGGCTAAGCTGAAGCTATTGGA	840
Qy	841	ATCCTGGCTGAGCTCTGACACAAATATGAGATGAGCAGCTTCACTGCTGCGCC	900
Db	841	ATCCTGGCTGAGCTCTGACACAAATATGAGATGAGCAGCTTCACTGCTGCGCC	900
Qy	901	GAGCAGTATGTACGGCGTCTCCAACTGCGCCAGGACTCCAACTATCTACTGCGCC	960
Db	901	GAGCAGTATGTACGGCGTCTCCAACTGCGCCAGGACTCCAACTATCTACTGCGCC	960
Qy	961	TCCAACTGCGGATGTACACAGATGTGCTGAGGCGATGGGTGTATGTGGAGCCCTC	1020
Db	961	TCCAACTGCGGATGTACACAGATGTGCTGAGGCGATGGGTGTATGTGGAGCCCTC	1020
Qy	1021	ACCAAGGCGGAGTCCAGAGGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGTC	1080
Db	1021	ACCAAGGCGGAGTCCAGAGGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGTC	1080

Qy	1081	CAGGTACAGATGCAAGTNTTGTATGAGGAACTTGTATGATGAGTCAAGATGAGTGTAGTGGAGC	1140
Db	1081	CAGGTACAGATGCAAGTNTTGTATGAGGAACTTGTATGATGAGTCAAGATGAGTGTAGTGGAGC	1140
Qy	1141	TGGCTTNGCCAGGAGTCTCGGGACAGGAAGCAGATTTTCCCTGATT	1188
Db	1141	TGGCTTNGCCAGGAGTCTCGGGACAGGAAGCAGATTTTCCCTGATT	1188

RESULT 2

BC003025 Homo sapiens, stomachin-like 2, clone MGC:4191 IMAGE:2821269, mRNA, complete cds. PRI 12-JUL-2001

ACCESSION BC003025.1 GI:12804332
KEYWORDS MGC.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1257)
Strausberg, R.
Direct Submission
Submitted (06-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/BTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalius, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott zuiderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 9 Row: 1 Column: 7
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 9652258.

FEATURES

source

Location/Qualifiers
1..1257
/organism="Homo sapiens"
/db_xref="LocusID:30968"
/db_xref="taxon:9606"
/clone="MGC:4191 IMAGE:2821269"
/tissue_type="Lung, small cell carcinoma"
/clone_lib="NIH_MGC_7"
/lab_host="DH10B-R"
/note="vector: pOTB7"
23..1093
/codon_start=1
/product="stomatin-like 2"
/protein_id="AAH03025.1"
/db_xref="GI:12804333"
/translation="MLARAAGTGCALLRGLSLASGRAPRRASGLPRNTVVLFPVPO
EAWVVRMGFRHLEPLNLIPLDRIYVOSLKELIVNPVQSVATLDNVTQLD
EVLVLRIMDPYKASYGVDEPYAVTQLAOTMRSELKLSLDKVRRESINASIVDA
INAAQDCWIGRLRYETKDIHVPVKESQMOMVEAPRRKRTVLESEGTRESINVA
EGKQAGILASEAKAEQINQAAGEASAVLAKAKAKAEIRILAAATONGDAASL
TVAEQIVSAFSLKAKDSNTILLPSNPQDVTSMAVQAMGVYCALTKAPVPGTDPDLSLGG

CDS

Query Match 2.0%; Score 7; DB 1; Length 16;
 Best Local Similarity 100.0%; Pred. No. 5.8;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 249 AKAKAKA 255
 Db 2 AKAKAKA 8

RESULTS
 US-08-346-849-25
 ; Sequence 25, Application US/08346849
 ; Patent No: 5670483
 ; GENERAL INFORMATION:
 ; APPLICANT: Zhang, Shuguang
 ; APPLICANT: Lockshin, Curtis
 ; APPLICANT: Rich, Alexander
 ; APPLICANT: Holmes, Todd
 ; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
 ; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
 ; NUMBER OF SEQUENCES: 64
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
 ; STREET: Two Millitia Drive
 ; CITY: Lexington
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02173-4799
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/346,849
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/973,326
 ; FILING DATE: 28 DECEMBER 1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Brook, David E.
 ; REGISTRATION NUMBER: 22,592
 ; REFERENCE/DOCKET NUMBER: MIT-6008
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 861-6240
 ; TELEFAX: (617) 861-9540
 ; INFORMATION FOR SEQ ID NO: 25:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 16 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; 346-849-25

SSRDVOGCTDASLDEELDRKMS"									
BASE COUNT	309 a	324 c	364 g	260 t					
ORIGIN									
Query Match	72.1%;	Score 857;	DB 9;	Length 1257;					
Best Local Similarity	99.8%;	Pred. No. 0;							
Matches 957;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;					
QY	140	GC	CGCGCTCTCTCTGGATTGCCCGAACACCGTGGTACTGTTGTCGCCGACGAGGAGG	199					
Db	99	GC	CGCGCTCTCTCTGGATTGCCCGAACACCGTGGTACTGTTGTCGCCGACGAGGAGG	158					
QY	200	CCT	GGGTGGAGCGAATGGCGGATTCACCGGATCCTGGAGCCCTGGTTTGAACATCC	259					
Db	159	CCT	GGGTGGAGCGAATGGCGGATTCACCGGATCCTGGAGCCCTGGTTTGAACATCC	218					
QY	260	TCAT	CCCTGTGTTAGACCGGATCGGATATGTCAGAGTCTCAAGGAAATTCATCAACG	319					
Db	219	TCAT	CCCTGTGTTAGACCGGATCGGATATGTCAGAGTCTCAAGGAAATTCATCAACG	278					
QY	320	TGCG	TGAGCAGTCGGCTGTGACTCTGCACAACTGTAACCTCTGCAATCGATGGAGTCCTTT	379					
Db	279	TGCG	TGAGCAGTCGGCTGTGACTCTGCACAACTGTAACCTCTGCAATCGATGGAGTCCTTT	338					
QY	380	ACCT	GCCATCAGTGGACCCCTACAGGCAAGCTACGGTGTGGAGACCCCTGAGTATGCCG	439					
Db	339	ACCT	GCCATCAGTGGACCCCTACAGGCAAGCTACGGTGTGGAGACCCCTGAGTATGCCG	398					
QY	440	TCAC	CCAGCTAGCTCAACCAACCATGATGATCAGAGCTCGGCAAACTCTCTTGGACAAG	499					
Db	399	TCAC	CCAGCTAGCTCAACCAACCATGATGATCAGAGCTCGGCAAACTCTCTTGGACAAG	458					
QY	500	TCCT	CCGGGAACGGAGTCCCTGAATGCCAGCATTTGGATGCCATCAACCAAGCTGCTG	559					
Db	459	TCCT	CCGGGAACGGAGTCCCTGAATGCCAGCATTTGGATGCCATCAACCAAGCTGCTG	518					
QY	560	ACTG	CTGGGTATCCGCTGCCCTNGTTATGAGATCAAGGATATCCATGTGCCACCCCGGG	619					
Db	519	ACTG	CTGGGTATCCGCTGCCCTNGTTATGAGATCAAGGATATCCATGTGCCACCCCGGG	578					
QY	620	TGAA	GAGTCTATGCAGATGCAGTGGAGGACGAGCGCGGCAAGACGGCCACAGTTCTAG	679					
Db	579	TGAA	GAGTCTATGCAGATGCAGTGGAGGACGAGCGCGGCAAGACGGCCACAGTTCTAG	638					
QY	680	AGTC	TGAGGGACCCGAGAGTGCGCCATCAATGTGGCAGAGGAGGAACAGAGCCCGAGA	739					
Db	639	AGTC	TGAGGGACCCGAGAGTGCGCCATCAATGTGGCAGAGGAGGAACAGAGCCCGAGA	698					
QY	740	TCCT	GGCTCCGAGCAGAAAAGCTCAACAGATAAATCAGGCAGCAGGAGAGGCCAGTG	799					
Db	699	TCCT	GGCTCCGAGCAGAAAAGCTCAACAGATAAATCAGGCAGCAGGAGAGGCCAGTG	758					
QY	800	CAGT	TTGGGGAAGGCCAAAGCTAAAGCTGAAGCTATTTCGAATCCTGGCTGCAGCTCTGA	859					
Db	759	CAGT	TTGGGGAAGGCCAAAGCTAAAGCTGAAGCTATTTCGAATCCTGGCTGCAGCTCTGA	818					
QY	860	CACA	ACATAATGGAGATGCAGCAGCTTCACTGACTGTGGCGGACGACATATGTCAGCGCT	919					
Db	819	CACA	ACATAATGGAGATGCAGCAGCTTCACTGACTGTGGCGGACGACATATGTCAGCGCT	878					
QY	920	TCCT	CAAACTGGCCAAAGGACTTCCAACACTATCTTACTTGGCCCTCCAACCCCTGGCGATGCA	979					
Db	879	TCCT	CAAACTGGCCAAAGGACTTCCAACACTATCTTACTTGGCCCTCCAACCCCTGGCGATGCA	938					
QY	980	CCAG	CATGGTGGCTCAGGCCATGGGTGTATATGAGGAGCCCTCACCAAAAGCCCGATGCCAG	1039					
Db	939	CCAG	CATGGTGGCTCAGGCCATGGGTGTATATGAGGAGCCCTCACCAAAAGCCCGATGCCAG	998					
QY	1040	GGAC	TCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGTCCAGGGTACAGATCAAGT	1098					
Db	999	GGAC	TCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGTCCAGGGTACAGATCAAGT	1057					

```
QY 260 TCATCCCTGTGTTAGACCGGATCCGATATGTGACAGTCTCAAGGAAATTTGTCATCAACG 319
Db 231 TCATCCCTGTGTTAGACCGGATCCGATATGTGACAGTCTCAAGGAAATTTGTCATCAACG 290
QY 320 TGCCTGAGCAGTGGCTGTGACTCTCACAATGTAACTCTGCAATCGATGGAGTCTCTTT 379
Db 291 TGCCTGAGCAGTGGCTGTGACTCTCACAATGTAACTCTGCAATCGATGGAGTCTCTTT 350
QY 380 ACCTGCGCATCATGACGCCCTTACAAAGCAAGCTACGGTGTGGAGACCCCTGAGTATCCG 439
Db 351 ACCTGCGCATCATGAGCCCTTACAAAGCAAGCTACGGTGTGGAGACCCCTGAGTATCCG 410
QY 440 TCACCCAGCTAGCTCAAAACACCATGATGATCAGAGCTCGGCAAACTCTCTNTGACAAAG 499
Db 411 TCACCCAGCTAGCTCAAAACACCATGATGATCAGAGCTCGGCAAACTCTCTGACAAAG 470
QY 500 TCTTCCCGGAACGGGAGTCCCTGAAATGCCAGCATTTGGATGCCATCAACCAAGCTGCTG 559
Db 471 TCTTCCCGGAACGGGAGTCCCTGAAATGCCAGCATTTGGATGCCATCAACCAAGCTGCTG 530
QY 560 ACTGCTGGGGTATCCGCTGCTCTGCTTATGAGATCAAGGATATCCATGTGCCACCCCGGG 619
Db 531 ACTGCTGGGGTATCCGCTGCTCTGCTTATGAGATCAAGGATATCCATGTGCCACCCCGGG 590
QY 620 TGAAGAGTCTATGACAGATGAGTGGAGGACGAGCGCGGAAACGGGCGCACAGTCTTAG 679
Db 591 TGAAGAGTCTATGACAGATGAGTGGAGGACGAGCGCGGAAACGGGCGCACAGTCTTAG 650
QY 680 AGTCTGAGGGACCCGAGAGTCCGCCATCAATGTGGCAGAGGGAAGAAACAGGCCCCAGA 739
Db 651 AGTCTGAGGGACCCGAGAGTCCGCCATCAATGTGGCAGAGGGAAGAAACAGGCCCCAGA 710
QY 740 TCTTGGCTCCGAAAGCAAGGCTGAACAGATATAATCAGGACGAGGAGAGGCCAGT 799
Db 711 TCTTGGCTCCGAAAGCAAGGCTGAACAGATATAATCAGGACGAGGAGAGGCCAGT 770
QY 800 CAGTTCTGGGCAAGGCCAAGCTTAAAGCTGAAGCTATTCGAATCCTGGCTGACGCTCTGA 859
Db 771 CAGTTCTGGGCAAGGCCAAGCTTAAAGCTGAAGCTATTCGAATCCTGGCTGACGCTCTGA 830
QY 860 CACAACATAATGAGATGACAGCTTCACTGACTGTGGCGGACGACATATGTCAGCGCT 919
Db 831 CACAACATAATGAGATGACAGCTTCACTGACTGTGGCGGACGACATATGTCAGCGCT 890
QY 920 TCTCCAAACTGGCCAAAGGACTCCAAACACTATCTCTACTGCCCTCCAAACCCCTGGCGATGCA 979
Db 891 TCTCCAAACTGGCCAAAGGACTCCAAACACTATCTCTACTGCCCTCCAAACCCCTGGCGATGCA 950
QY 980 CCAGCATGGTGGCTCAGGCCATGGGTGTATATGGAGCCCTCACAAAGCCCGCAGTGCAG 1039
Db 951 CCAGCATGGTGGCTCAGGCCATGGGTGTATATGGAGCCCTCACAAAGCCCGCAGTGCAG 1010
QY 1040 GGACTCCAGACTCACTCTCAGTGGGAGCAGCAGATGTCCAGGTCACAGATCAAGT 1098
Db 1011 GGACTCCAGACTCACTCTCAGTGGGAGCAGCAGATGTCCAGGTCACAGATCAAGT 1069

RESULT 4
LOCUS BC014990 1262 bp mRNA linear PRI 29-OCT-2001
DEFINITION Homo sapiens, stomatin-like 2, clone MGC:23225 IMAGE:4907745, mRNA, complete cds.
ACCESSION BC014990
VERSION BC014990.1 GI:15929069
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1262)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (01-OCT-2001) National Institutes of Health, Mammalian
```

REMARK
COMMENT

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garlund, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lilisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 34 Row: p Column: 12
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES
source

Location/Qualifiers
1. .1262
/organism="Homo sapiens"
/db_xref="LocusID:30968"
/db_xref="taxon:9606"
/clone="MGC:23225 IMAGE:4907745"
/tissue_type="Pancreas, epithelioid carcinoma"
/clone_lib="NIH_MGC_42"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
33. .1103
/codon_start=1
/product="stomatin-like 2"
/protein_id="AAH14990.1"
/db_xref="GI:15929070"
/translation="MLAARAAGTGALLRGLSLASRAPRRASGGLPRNTVVLFPVPOO
EAWVERMGRFHRIEPLVDRIYVQSLKEIVINVPESQAVTLDNVTLOID
GVLVLRIMDPYKASYGVDEPYAVTOLAOTMRSELCKSLDKVFRRESLSAIVDA
INQAAOCWGRICLRYEIKDIHVPVRKESMOMOVEAEKRRKRVLESEGTRESAIVDA
EGKKAQILASEAEAEQINQAGEASAVLAKAKAEAIRILAAALTOHNGDAASL
TVAEQIVSAFSKLUAKDSNTILLPSNFGDVTSMVAQAMGVYGTALTAPVPGTDSLSSG
SSRDVQGTDAISLDELDRVKMS"

CDS

BASE COUNT 305 a 326 c 367 g 264 t
ORIGIN

Query Match 72.1%; Score 857; DB 9; Length 1262;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 140 GCCGCGCTCTCTGGATTGCCCGGAAACACCGTGGTACTTCTGTCGCCGACGAGGAGG 199
Db 109 GCCGCGCTCTCTGGATTGCCCGGAAACACCGTGGTACTTCTGTCGCCGACGAGGAGG 168
QY 200 CCTGGGTGGTGGAGCGAATGGCGGATTCCACCGGATCCTGGAGCCCTGGTTTGAACATCC 259
Db 169 CCTGGGTGGTGGAGCGAATGGCGGATTCCACCGGATCCTGGAGCCCTGGTTTGAACATCC 228
QY 260 TCATCCCTGTGTTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAAATGTCATCAACG 319
Db 229 TCATCCCTGTGTTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAAATGTCATCAACG 288
QY 320 TGCTGAGCAGTCCGGCTGTGACTCTCGACAAATGTAACTCTGCAAAATCGATGAGTCCCTTT 379
Db 289 TGCTGAGCAGTCCGGCTGTGACTCTCGACAAATGTAACTCTGCAAAATCGATGAGTCCCTTT 348

; TITLE: N/A
 ; PUBLICATION INFORMATION:
 ; DOCUMENT NUMBER: PCT/US92/08069
 ; FILING DATE: 14-AUG-1993
 ; US-08-152-488-3

Query Match 2.0%; Score 7; DB 1; Length 29;
 Best Local Similarity 100.0%; Pred. No. 9.8;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 AKAKAKA 255
 Db 4 AKAKAKA 10

RESULT 7
 US-08-152-488-3
 ; Sequence 3, Application US/08152488
 ; Patent No. 5534619
 ; GENERAL INFORMATION:
 ; APPLICANT: Wakefield, Thomas W.
 ; APPLICANT: Andrews, Philip C.
 ; APPLICANT: Stanley, James C.
 ; TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
 ; TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
 ; TITLE OF INVENTION: ANTICOAGULATION REVERSAL
 ; NUMBER OF SEQUENCES: 13
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Benita J. Rohm, Esq.
 ; STREET: 512 Springfield Avenue
 ; CITY: Cranford
 ; STATE: New Jersey
 ; COUNTRY: United States of America
 ; ZIP: 07016-1811
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC Compatible
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: WordPerfect 6; ASCII (DOS)Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/152,488
 ; FILING DATE: 12-NOV-1993
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US92/08069
 ; FILING DATE: 14-AUG-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Rohm, Benita J.
 ; REGISTRATION NUMBER: 28,664
 ; REFERENCE/DOCKET NUMBER: RM-7WG
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 908-276-3344
 ; TELEFAX: 908-276-5543
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 29 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: N/A
 ; TOPOLOGY: N/A
 ; MOLECULE TYPE: peptide
 ; ORIGINAL SOURCE:
 ; ORGANISM: N/A
 ; PUBLICATION INFORMATION:
 ; AUTHORS: N/A


```
QY 380 ACCTGGCATCATGACCCCTTACAAGGCAAGCTACGGTGTGGAGGACCTGAGTATGCGG 439
Db 349 ACCTGGCATCATGACCCCTTACAAGGCAAGCTACGGTGTGGAGGACCTGAGTATGCGG 408
QY 440 TCACCCAGCTAGCTCAAAACAACCATGAGATCAGAGCTCGGCAAACTCTCTNTGGACAAG 499
Db 409 TCACCCAGCTAGCTCAAAACAACCATGAGATCAGAGCTCGGCAAACTCTCTNTGGACAAG 468
QY 500 TCTTCCGGGAACGGGAGTCCCTGAATGCCAGCATTTGTGGATCCCATCAACCAAGTGTGTG 559
Db 469 TCTTCCGGGAACGGGAGTCCCTGAATGCCAGCATTTGTGGATCCCATCAACCAAGTGTGTG 528
QY 560 ACTGCTGGGGTATCCGCTGCCCTTCATGATGATCAAGATATCCATGTCGCCACCCCGGG 619
Db 529 ACTGCTGGGGTATCCGCTGCCCTTCATGATGATCAAGATATCCATGTCGCCACCCCGGG 588
QY 620 TCAAAAGATCTATGAGATGCGAGTGGAGGAGAGCGCGGGAAGAGGAGGAGGAGGAGGAG 679
Db 589 TCAAAAGATCTATGAGATGCGAGTGGAGGAGAGCGCGGGAAGAGGAGGAGGAGGAGGAG 648
QY 680 AGTCTGAGGGGACCGGAGAGTGGGCGCATCAATGTGGCAGAGGAGGAGGAGGAGGAGGAG 739
Db 649 AGTCTGAGGGGACCGGAGAGTGGGCGCATCAATGTGGCAGAGGAGGAGGAGGAGGAGGAG 708
QY 740 TCCTGGCTCCGAAGCAAGAGGCTGAACAGATTAATCAGGAGGAGGAGGAGGAGGAGGAG 799
Db 709 TCCTGGCTCCGAAGCAAGAGGCTGAACAGATTAATCAGGAGGAGGAGGAGGAGGAGGAG 768
QY 800 CAGTTCTGCGAAGGCCAAGGCTAAAGCTGAAGCTTATTCGAATCTCTGCTGCGAGCTCTGA 859
Db 769 CAGTTCTGCGAAGGCCAAGGCTAAAGCTTATTCGAATCTCTGCTGCGAGCTCTGA 828
QY 860 CACAACATAATGGAGATGAGCAGGCTTCACTGACTGTGGCCGAGCAGTATGTGACGCGGT 919
Db 829 CACAACATAATGGAGATGAGCAGGCTTCACTGACTGTGGCCGAGCAGTATGTGACGCGGT 888
QY 920 TCTCAAACTGCCAAGGAGCTCAACACTATCTACTGCTCCCTCAACCTGCGGAGTGTCA 979
Db 889 TCTCAAACTGCCAAGGAGCTCAACACTATCTACTGCTCCCTCAACCTGCGGAGTGTCA 948
QY 980 CCAGCATGTGCTCAGGCCATGGGTGTATATGGAGCCCTCACCAGGAGGAGGAGGAGGAG 1039
Db 949 CCAGCATGTGCTCAGGCCATGGGTGTATATGGAGCCCTCACCAGGAGGAGGAGGAGGAG 1008
QY 1040 GGACTCCAGACTCACTCCAGTGGGAGGAGCAGAGATGTCAGGAGTACAGATGCAAGT 1098
Db 1009 GGACTCCAGACTCACTCCAGTGGGAGGAGCAGAGATGTCAGGAGTACAGATGCAAGT 1067
```

```
RESULT 5
BC002442
LOCUS Homo sapiens, stomatin-like 2, clone MGC:1179 IMAGE:3346384, mRNA,
DEFINITION complete cds.
ACCESSION BC002442
VERSION BC002442.1 GI:12803254
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1278)
Direct Submission
Submitted (05-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
```

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hgrl.nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, O.L., Masiello, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, J.J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,
Tiongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
Zhang, L.-H., and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 5 Row: m Column: 8
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 9652258.

FEATURES

Location/Qualifiers
1..1278
/organism="Homo sapiens"
/db_xref="LocusID:30968"
/db_xref="taxon:9606"
/clone="MGC:1179 IMAGE:3346384"
/tissue_type="skin, melanotic melanoma."
/clone_lib="NIH_MGC_20"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
47..1117
/codon_start=1
/product="stomatin-like 2"
/protein_id="AAH02442.1"
/db_xref="GI:12803255"
/translation="MLAARAGTCALLRGLSLASGRAPRRASSGLPNTVVLFPQQ
EAWYVERGFRHLEPGLNLIPLVDLRITVOSLUKEIVINVPESAVFLDNVTLQID
GVLYLRINDPYKASYDEVEDPYATVLAQTSLRSLDKVRESLRESLRESAINVA
INQADGWIRCLRYEIKDIHVPPRKESQMOMVEARRKRVLESEGTRESAINVA
EGKKQAOILASEAEKAEQINOAGASAVLAKAKAEAIRILAAALTOHNGDAASL
TVAEQVVSASFSLKADSNITILLPSNPGDVTSMVAQMGVYGTALTRAPVGTPTDLSLSS
SSRDVGTGTDASLDELDRVMS"
BASE COUNT 309 a 330 c 373 g 266 t
ORIGIN

Query Match 72.1%; Score 857; DB 9; Length 1278;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 140 GCCGCGCTCTCTGATTGCCCGAAGACCGTGTGTTCTGTCGCGCAGCAGGAG 199
Db 123 GCCGCGCTCTCTGATTGCCCGAAGACCGTGTGTTCTGTCGCGCAGCAGGAG 182
QY 200 CCTGGGTGTGGAGCAATGGCGGATCCACGGGATCTCGGAGCCTGTTTGAACATCC 259
Db 183 CCTGGGTGTGGAGCAATGGCGGATCCACGGGATCTCGGAGCCTGTTTGAACATCC 242
QY 260 TCATCCCTGTGTAGACCGGATCCGATATGTGACAGTCTCAAGGAAATGTCATCAACG 319
Db 243 TCATCCCTGTGTAGACCGGATCCGATATGTGACAGTCTCAAGGAAATGTCATCAACG 302
QY 320 TGCCTGAGCAGTCGGCTGTGACTCTCGAACAATGTAACCTGCAAAATCGATGGAGTCTTT 379
Db 303 TGCCTGAGCAGTCGGCTGTGACTCTCGAACAATGTAACCTGCAAAATCGATGGAGTCTTT 362
QY 380 ACCTGGCATCATGAGACCTTACAAGGCAAGCTACGGTGTGGAGGAGCCTGAGTATGCGG 439
Db 363 ACCTGGCATCATGAGACCTTACAAGGCAAGCTACGGTGTGGAGGAGCCTGAGTATGCGG 422
QY 440 TCACCCAGCTAGCTCAAAACAACCATGAGATCAGAGCTCGGCAAACTCTCTNTGGACAAG 499
Db 423 TCACCCAGCTAGCTCAAAACAACCATGAGATCAGAGCTCGGCAAACTCTCTNTGGACAAG 482

```
Qy 500 TCTTCGGGAAGCGGAGTCCCTGAATGCCAGATGTGTGGATGCCATCAACCAAGCTGCTG 559
Db 483 TCTTCGGGAAGCGGAGTCCCTGAATGCCAGATGTGTGGATGCCATCAACCAAGCTGCTG 542
Qy 560 ACTGCTGGGGTATCCGCTGCCTTCGCTTATGAGATCAAGGATATCCATGTGCCACCCCGGG 619
Db 543 ACTGCTGGGGTATCCGCTGCCTTCGCTTATGAGATCAAGGATATCCATGTGCCACCCCGGG 602
Qy 620 TGAAGAGTCTATGATGATGATGAGGAGGAGCGGCGGGAAGCGGCCACAGTTCCTAG 679
Db 603 TGAAGAGTCTATGATGATGAGGAGGAGCGGCGGGAAGCGGCCACAGTTCCTAG 662
Qy 680 AGTCTGAGGGACCCGAGAGTCGGCCATCAATGTGGCAGAAGGAAGAACAGGCCCCAGA 739
Db 663 AGTCTGAGGGACCCGAGAGTCGGCCATCAATGTGGCAGAAGGAAGAACAGGCCCCAGA 722
Qy 740 TCCTGGCTCCGAAGCAAGAGGCTGAACAGATTAATCAGGCAGCAGAGAGGCCAGTG 799
Db 723 TCCTGGCTCCGAAGCAAGAGGCTGAACAGATTAATCAGGCAGCAGAGAGGCCAGTG 782
Qy 800 CAGTTCTGGCGAAGCCCAAGGCTAAGCTGAAGCTATTTCGAATCCTGGCTGCAGCTCTGA 859
Db 793 CAGTTCTGGCGAAGCCCAAGGCTAAGCTGAAGCTATTTCGAATCCTGGCTGCAGCTCTGA 842
Qy 860 CACAACATAATGAGATGCAGCAGCTTCACTGACTGTGGCCGAGCAGTATGTCAAGCGGT 919
Db 843 CACAACATAATGAGATGCAGCAGCTTCACTGACTGTGGCCGAGCAGTATGTCAAGCGGT 902
Qy 920 TCTCCAAACTGCCAAGACATCCACACTATCCTACTGCCCTCCAAACCTGCCGATGTCA 979
Db 903 TCTCCAAACTGCCAAGACATCCACACTATCCTACTGCCCTCCAAACCTGCCGATGTCA 962
Qy 980 CCAGCATGGTGGCTCAGGCCATGGGTGTATATGAGGCCCTCACCAAGCCCGCAGTGCAG 1039
Db 963 CCAGCATGGTGGCTCAGGCCATGGGTGTATATGAGGCCCTCACCAAGCCCGCAGTGCAG 1022
Qy 1040 GGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGT 1098
Db 1023 GGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGT 1081

RESULT 6
AF190167 AF190167 1303 bp mRNA linear PRI 13-MAR-2000
DEFINITION Homo sapiens membrane associated protein SLP-2 (HUSLP2) mRNA,
complete cds.
ACCESSION AF190167
VERSION AF190167.1 GI:6456117
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1303)
Wang, Y. and Morrow, J.S.
Identification and characterization of human SLP-2, a novel
homologue of stomatin (band 7.2b) present in erythrocytes and other
tissues
J. Biol. Chem. 275 (11), 8062-8071 (2000)
MEDLINE 2017914
REFERENCE 2 (bases 1 to 1303)
AUTHORS Wang, Y. and Morrow, J.S.
TITLE Direct Submission
JOURNAL Submitted (25-SEP-1999) Pathology, Yale Medical School, 310-Cedar
Street, New Haven, CT 06510, USA
FEATURES
source Location/Qualifiers
1..1303
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/tissue_type="heart muscle"
1..1303
/gene="HUSLP2"
gene
```

```
CDS
64..1134
/gene="HUSLP2"
/note="stomatin-like protein 2; widely distributed
peripheral membrane protein; similar to human erythrocyte
stomatin and MEC 1 of Caenorhabditis elegans; thought to
be involved in mechanoreception or lipid domain
organization"
/codon_start=1
/product="membrane associated protein SLP-2"
/protein_id="AAF09142.1"
/db_xref="GI:6456118"
/translation="MLAARAGTGCALLRSLGSLASGRAPRRASSGLPRNTVVLFVPOQ
EAWVERMGRFHRILEPGLNILIPVLDRIYVQSLKEIVINPEQSAVTLDNVTLQID
GVLVLRIMDPYKASGYGVDEYATQLAQTMRSELGKLSLDKVFRRERESINASTIDA
INQADWGIKRLRYEIKDHPVPRKESMOMOVEABRRKRATVLESEGTRESAINVA
EGKQAOILASEAEKAEQINQAGEASAVLAKAKAKAEAIRLLAAALTOHNGDAASL
TVABQYVSFASKLAKDSNTILLPSNPGDVTSWAQAMGVGYALTAKVPVPGTDSLSSG
SSRDVQGTDAASLDELDRVXMS"
217
/misc_feature
/gene="HUSLP2"
/note="alternative translation initiation site"
391
/misc_feature
/gene="HUSLP2"
/note="alternative translation initiation site"
BASE COUNT 315 a 338 c 378 g 272 t
ORIGIN

Query Match 72.1%; Score 857; DB 9; Length 1303;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 140 GCCGGCCCTCTCTGGATTGCCCGCAACACCGTGGTACTGTTCGTGCGCAGCAGGAGG 199
Db 140 GCCGGCCCTCTCTGGATTGCCCGCAACACCGTGGTACTGTTCGTGCGCAGCAGGAGG 199
Qy 200 CCTGGGTGGTGAGGGAATGGCCGATTCACCCGGATCCTGGAGCCTGGTTTGAACATCC 259
Db 200 CCTGGGTGGTGAGGGAATGGCCGATTCACCCGGATCCTGGAGCCTGGTTTGAACATCC 259
Qy 260 TCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAAATGTGTCAACG 319
Db 260 TCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAAATGTGTCAACG 319
Qy 320 TGCCTGAGCAGTCGCTGTGACTCTCGACAATGTAACTCTGCMAATCGATGAGTCCCTTT 379
Db 320 TGCCTGAGCAGTCGCTGTGACTCTCGACAATGTAACTCTGCMAATCGATGAGTCCCTTT 379
Qy 380 ACCTCGGCATCATGACCCCTTACAAGGCAAGCTACGGTGTGGAGGACCCCTGAGTATGCCG 439
Db 380 ACCTCGGCATCATGACCCCTTACAAGGCAAGCTACGGTGTGGAGGACCCCTGAGTATGCCG 439
Qy 440 TCACCCAGCTAGCTCAAAACACCATGAGATCAGAGCTCGGCAAACTCTCTTGGACAAG 499
Db 440 TCACCCAGCTAGCTCAAAACACCATGAGATCAGAGCTCGGCAAACTCTCTTGGACAAG 499
Qy 500 TCTTCGGGGAACGGGAGTCCCTGGAATGCCAGCATTTGTGGATGCCATCAACCAAGTGTGTG 559
Db 500 TCTTCGGGGAACGGGAGTCCCTGGAATGCCAGCATTTGTGGATGCCATCAACCAAGTGTGTG 559
Qy 560 ACTGCTGGGGTATCCGCTGCCTTCGCTTATGAGATCAAGGATATCCATGTGCCACCCCGGG 619
Db 560 ACTGCTGGGGTATCCGCTGCCTTCGCTTATGAGATCAAGGATATCCATGTGCCACCCCGGG 619
Qy 620 TGAAGAGTCTATGAGATGAGGAGGAGCGGCGGGAAGCGGCCACAGTTCCTAG 679
Db 620 TGAAGAGTCTATGAGATGAGGAGGAGCGGCGGGAAGCGGCCACAGTTCCTAG 679
Qy 680 AGTCTGAGGGACCCGAGAGTCGGCCATCAATGTGGCAGAAGGAAGAACAGGCCCCAGA 739
Db 680 AGTCTGAGGGACCCGAGAGTCGGCCATCAATGTGGCAGAAGGAAGAACAGGCCCCAGA 739
Qy 740 TCCTGGCTCCGAAGCAAGAGGCTGAACAGATTAATCAGGCAGCAGAGAGGCCAGTG 799
```

```
Db 740 TCCTGGCTCCGAAGCAAGAGGCTGAACAGATAAATCAGGCAGCAGGAGGCGCATG 799
Qy 800 CAGTTCTGGCAGCCAGGCTAAAGCTGAAGCTATTTCGAATCTGGCTGCAGCTCTGA 859
Db 800 CAGTTCTGGCAGCCAGGCTAAAGCTGAAGCTATTTCGAATCTGGCTGCAGCTCTGA 859
Qy 860 CACAACATAAATGGAGATGCAGAGCTTCACTGACTGTGGCCGAGCAGTATGTCAGCGGT 919
Db 860 CACAACATAAATGGAGATGCAGAGCTTCACTGACTGTGGCCGAGCAGTATGTCAGCGGT 919
Qy 920 TCTCCAAACTGCCAAGACATCCACACATATCTACTGCTCCCTCCAAACCTGGCGATGCA 979
Db 920 TCTCCAAACTGCCAAGACATCCACACATATCTACTGCTCCCTCCAAACCTGGCGATGCA 979
Qy 980 CCAGCATGGTGGCTCAGGCCATGGGTGTATATGGAGCCCTCACCAGCCAGTGCAG 1039
Db 980 CCAGCATGGTGGCTCAGGCCATGGGTGTATATGGAGCCCTCACCAGCCAGTGCAG 1039
Qy 1040 GGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGTCCAGGTTACAGATGCAAGT 1098
Db 1040 GGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGTCCAGGTTACAGATGCAAGT 1098

RESULT 7
AK027405 1244 bp mRNA linear PRI 15-MAY-2001
LOCUS Homo sapiens cDNA FLJ14499 fis, clone NT2RM1000080, weakly similar
DEFINITION to UNC-1 PROTEIN.
ACCESSION AK027405
VERSION AK027405.1 GI:14042059
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens teratocarcinoma cell_line:NT2 cDNA to mRNA,
clone_lib:NT2RM1 clone:NT2RM1000080.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
Wagatsuma,M., Hosoi,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
Nakamura,Y., Nagahara,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
NEO human cDNA sequencing project
Unpublished
2 (bases 1 to 1244)
Isogai,T. and Otsuki,T.
Direct Submission
JOURNAL Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)
COMMENT NEO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology; cDNA library construction;
5'- & 3'-end one pass sequencing and clone selection; Helix
Research Institute (supported by Japan Key Technology Center etc.)
and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
Location/Qualifiers
source 1..1244
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RM1000080"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RM1"
/note="cloning vector: pUC19FL3-mRNA from uninduced NT2
neuronal precursor cells."
35..1105
/note="unnamed protein product"
/codon_start=1
CDS
```

```
/protein_id="BAB55091.1"
/db_xref="gi:14042060"
/translation="MLAARAGTGALLIRGSLASGRAPRRASSGLPRTNVLFVPQO
EAWVERMRGPHRILEPGLNLIIPVLDRIYVQSLKEIVINVPQSAVTLDNVTLOID
GVLVLRIMDPYKASYGVVEDPEYATOLAQTMRSELKLSLDKVPRESLNASLIDA
INQADCWGIRCLYEIKDIIHVPVRKESMQVEARRKRATVLESEGTRESAINWA
EGKQAQILASEAKEDINOAGEASAVLAKAKAEAIRILAAALTOHNGDAAASL
TVABQYVSASFSLKARDSNTILLPSNPGDVTSMVAQMGVYGALTAPYPGTSDLSLSS
SSRDVQGTGTDASLDELDRVMS"
BASE COUNT 284 a 327 c 368 g 265 t
ORIGIN
Query Match 67.8%; Score 806; DB 9; Length 1244;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 956; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 140 GCCGGCCTCTCTCGATTGCCCGGAACACCGTGGTACTGTTCGTGCGCAGCAGGAGG 199
Db 111 GCCGGCCTCTCTCGATTGCCCGGAACACCGTGGTACTGTTCGTGCGCAGCAGGAGG 170
Qy 200 CCTGGGTGGTGGAGCAATGGGCCGATTCCACCGGATCCTCGAGGCTGCTTTGAACATCC 259
Db 171 CCTGGGTGGTGGAGCAATGGGCCGATTCCACCGGATCCTCGAGGCTGCTTTGAACATCC 230
Qy 260 TCATCCCTGTGTTAGACCCGATCCGATATGTGCAGAGTCTCAAGGAAATTTGTCATCAACG 319
Db 231 TCATCCCTGTGTTAGACCCGATCCGATATGTGCAGAGTCTCAAGGAAATTTGTCATCAACG 290
Qy 320 TGCCTGAGCAGTCCGGCTGTGACTCTCGACATGTACTCTGCAATTCGATGAGTCTTT 379
Db 291 TGCCTGAGCAGTCCGGCTGTGACTCTCGACATGTACTCTGCAATTCGATGAGTCTTT 350
Qy 380 ACCTCGGCATCATGACCCCTTACAAGGCAAGCTACGGTGTGGAGGACCTGAGTATGCCG 439
Db 351 ACCTCGGCATCATGACCCCTTACAAGGCAAGCTACGGTGTGGAGGACCTGAGTATGCCG 410
Qy 440 TCACCCAGCTAGCTCAAAACACCATGAGATCAGAGCTCGGCAAACTCTCTNTGGACAAAG 499
Db 411 TCACCCAGCTAGCTCAAAACACCATGAGATCAGAGCTCGGCAAACTCTCTCTGGACAAAG 470
Qy 500 TCTTCCGGGAAGGGAGTCCCTGAATGCCAGTATGTGGATGCCATCAACCAAGCTGTG 559
Db 471 TCTTCCGGGAAGGGAGTCCCTGAATGCCAGTATGTGGATGCCATCAACCAAGCTGTG 530
Qy 560 ACTGCTGGGGTATCCGCTGCCTTCGTTATGAGATCAAGGATATCCATGTGCCACCCCGG 619
Db 531 ACTGCTGGGGTATCCGCTGCCTTCGTTATGAGATCAAGGATATCCATGTGCCACCCCGG 590
Qy 620 TGAAGAGTCTATGCAGATGCAGGTGGAGGAGAGCGGCGGAAACGGCCACAGTTCTAG 679
Db 591 TGAAGAGTCTATGCAGATGCAGGTGGAGGAGAGCGGCGGAAACGGCCACAGTTCTAG 650
Qy 680 AGTCTGAGGGACCCGAGAGTCCGCATCAATGTGGCAGAAAGGAGAAACAGGCCACAG 739
Db 651 AGTCTGAGGGACCCGAGAGTCCGCATCAATGTGGCAGAAAGGAGAAACAGGCCACAG 710
Qy 740 TCCTGCGCTCCCAAGCAGAAAAAGGCTGAACAGATAAATCAGGCAGCAGGAGGCGCAGTG 799
Db 711 TCCTGCGCTCCCAAGCAGAAAAAGGCTGAACAGATAAATCAGGCAGCAGGAGGCGCAGTG 770
Qy 800 CAGTTCTGCGGAAGGCCAAGGCTAAAGCTGAAGCTATTTCGAATCTGGCTGCAGCTCTGA 859
Db 771 CAGTTCTGCGGAAGGCCAAGGCTAAAGCTGAAGCTATTTCGAATCTGGCTGCAGCTCTGA 830
Qy 860 CACAACATAAATGGAGATGCAGAGCTTCACTGACTGTGGCCGAGCAGTATGTCAGCGGT 919
Db 831 CACAACATAAATGGAGATGCAGAGCTTCACTGACTGTGGCCGAGCAGTATGTCAGCGGT 890
Qy 920 TCTCCAAACTGCCAAGGACTCCAACTATCTACTGCTCCCTCCAAACCTGGCGATGCA 979
Db 891 TCTCCAAACTGCCAAGGACTCCAACTATCTACTGCTCCCTCCAAACCTGGCGATGCA 950
```

Qy 980 CCAGCATGGTGGCTCAGCCCATGGGTGTATATGAGAGCCCTCACCAAGCCCAAGTGCAG 1039
Db 951 CCAGCATGGTGGCTCAGCCCATGGGTGTATATGAGAGCCCTCACCAAGCCCAAGTGCAG 1010
Qy 1040 GGACTCCAGACTCACTCTCCAGTGGGAGCAGACAGAGATGTCAGGGTACAGATGCAAGT 1098
Db 1011 GGACTCCAGACTCACTCTCCAGTGGGAGCAGACAGAGATGTCAGGGTACAGATGCAAGT 1069

RESULT 8
LOCUS BC010152 1281 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, stomatin-like 2, clone MGC:19715 IMAGE:3534656, mRNA, complete cds.
ACCESSION BC010152
VERSION BC010152.1 GI:14603402
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1281)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (02-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapps-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ruben Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hgrl.nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J., Zhang, L.-H., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A., and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 27 Row: d Column: 1
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES

source

1. .1281
/organism="Homo sapiens"
/db_xref="LocusID:30968"
/db_xref="taxon:9606"
/clone="MGC:19715 IMAGE:3534656"
/tissue_type="Lung, small cell carcinoma"
/clone_lib="NIH_MGC_7"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
44. .1114
/codon_start=1
/product="stomatin-like 2"
/protein_id="AAH10152.1"
/db_xref="GI:14603403"
/translation="MLAARAGTGALLRGLSLASGRPRRASSGLPRNTVFLVPOQ
EAWVVRMGFRHRLPLGLNLIPLVDIRYVQSLKEIVINVPQSAVTLDNVTLOID
GVLYLRMDPYKDEYDEYVAPOTPMRSELGKLSLDKVFRESLSINAVDA
INQADWCGRICLYEIKDHPVPRVKESQMOMVEARRKRAIVLESEGTRESAINVA
EGKKQAILASEAKEAEQINQAAGEASVLA KAKAEATRIILAAULTQHNGDAASL

CDS

BASE COUNT 315 a 330 c 372 g 264 t
ORIGIN
TVAEQVVSFAFSKLAKDSNTILLPSNPGDVTSMVAQAMGVYGALTAKPVPCTPDSLSG
SSRDVQGTASLDEELDVRKMS"
Query Match 67.8%; Score 806; DB 9; Length 1281;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 956; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 140 GCCGGGCTCTCTGGATTGCCCGAACAACACCGTGTCTACTGTTGTCGGCAGCAGGAGG 199
Db 120 GCCCGGCTCTCTGGATTGCCCGAACAACACCGTGTCTACTGTTGTCGGCAGCAGGAGG 179
Qy 200 CCTGGGTGGTGGAGGAATGGCGGATTCCACCGGATCTCGAGCCTCTGGAGCCTGTTTGAACATCC 259
Db 180 CCTGGGTGGTGGAGGAATGGCGGATTCCACCGGATCTCGAGCCTCTGGAGCCTGTTTGAACATCC 239
Qy 260 TCATCCCTGTCTTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAAATTTGTCATCAACG 319
Db 240 TCATCCCTGTCTTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAAATTTGTCATCAACG 299
Qy 320 TGCCTGAGCAGTCCGCTGTGACTCTCGACAATGTAACCTCTGCAAAATCGATGGAGTCTCTTT 379
Db 300 TGCCTGAGCAGTCCGCTGTGACTCTCGACAATGTAACCTCTGCAAAATCGATGGAGTCTCTTT 359
Qy 380 ACCTGCCGATCATGACCCCTTACAGGCAAGCTACGGTGTGGAGGACCTCGATGATGCGG 439
Db 360 ACCTGCCGATCATGACCCCTTACAGGCAAGCTACGGTGTGGAGGACCTCGATGATGCGG 419
Qy 440 TCACCCAGCTAGCTCAAAACACCATGAGATCAGAGCTCGGCAAACTCTCTNTGGACAAG 499
Db 420 TCACCCAGCTAGCTCAAAACACCATGAGATCAGAGCTCGGCAAACTCTCTNTGGACAAG 479
Qy 500 TCTTCCGGGAACGGAGTCCCTGATGCCAGCATTTGGATGCCATCAACCAAGCTGCTG 559
Db 480 TCTTCCGGGAACGGAGTCCCTGATGCCAGCATTTGGATGCCATCAACCAAGCTGCTG 539
Qy 560 ACTGCTGGGGTATCCGCTGCTNCGTTATGAGATCAAGGATATCCATGTGCACCCCGGG 619
Db 540 ACTGCTGGGGTATCCGCTGCTNCGTTATGAGATCAAGGATATCCATGTGCACCCCGGG 599
Qy 620 TGAAGAGTCTATCGAGATCGAGTGGAGGAGCGCGGAAACAGGCCACAGTCTCTAG 679
Db 600 TGAAGAGTCTATCGAGATCGAGTGGAGGAGCGCGGAAACAGGCCACAGTCTCTAG 659
Qy 680 AGTCTGAGGGACCCGAGTCCGCCATCAATGTGGCAGAGGGAAGAAACAGGCCCA 739
Db 660 AGTCTGAGGGACCCGAGTCCGCCATCAATGTGGCAGAGGGAAGAAACAGGCCCA 719
Qy 740 TCCTGGCTCCGAAGCAGAAAGGCTGAACAGATAAATCAGGCAGCAGGAGGCGCAGTG 799
Db 720 TCCTGGCTCCGAAGCAGAAAGGCTGAACAGATAAATCAGGCAGCAGGAGGCGCAGTG 779
Qy 800 CAGTTCTGGCGAAGCCCAAGGCTAAAGCTGAAGCTATTCGAATCTTGCTGCTGAGTCTGA 859
Db 780 CAGTTCTGGCGAAGCCCAAGGCTAAAGCTGAAGCTATTCGAATCTTGCTGCTGAGTCTGA 839
Qy 860 CACAACATAATGGAGATGCACAGCTTCACCTGCTGCTGCGCAGCAGTATGTCACCGGCT 919
Db 840 CACAACATAATGGAGATGCACAGCTTCACCTGCTGCTGCGCAGCAGTATGTCACCGGCT 899
Qy 920 TCTCCAACTGGCCAAAGGACTCCCAACACTATCTACTGCCCTCCAAACCTCGCGGATGCA 979
Db 900 TCTCCAACTGGCCAAAGGACTCCCAACACTATCTACTGCCCTCCAAACCTCGCGGATGCA 959
Qy 980 CCAGCATGGTGGCTCAGGCCATGGGTGTATATGAGAGCCCTCACCAAGCCCAAGTGCAG 1039
Db 960 CCAGCATGGTGGCTCAGGCCATGGGTGTATATGAGAGCCCTCACCAAGCCCAAGTGCAG 1019
Qy 1040 GGACTCCAGACTCACTCTCCAGTGGGAGCAGACAGAGATGTCAGGGTACAGATGCAAGT 1098
Db 1020 GGACTCCAGACTCACTCTCCAGTGGGAGCAGACAGAGATGTCAGGGTACAGATGCAAGT 1078

RESULT 9
AX339946
LOCUS AX339946 457 bp DNA linear PAT 10-JAN-2002
DEFINITION Sequence 193 from Patent WO0196388.
ACCESSION AX339946
VERSION AX339946.1 GI:18135927
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Jiang, Y., Harlocker, S.L. and Secrist, H.
TITLE Compositions and methods for the therapy and diagnosis of colon cancer
JOURNAL Patent: WO 0196388-A 193 20-DEC-2001;
CORIXA CORPORATION (US)
FEATURES
source Location/Qualifiers
1..457
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 122 a 123 c 137 g 73 t 2 others
ORIGIN
Query Match 38.2%; Score 454; DB 6; Length 457;
Best Local Similarity 100.0%; Pred. No. 2e-262;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 633 GCAGATGCGAGTGGAGGAGCGGCGGAAACGGCCACAGTCTTAGAGTCTGAGGGGAC 692
Db 1 GCAGATGCGAGTGGAGGAGCGGCGGAAACGGCCACAGTCTTAGAGTCTGAGGGGAC 60
QY 693 CCGAGAGTGGCCATCAATGTGGCAGAGGGAAGAAACAGGCCAGATCCTGGCCCTCGCA 752
Db 61 CCGAGAGTGGCCATCAATGTGGCAGAGGGAAGAAACAGGCCAGATCCTGGCCCTCGCA 120
QY 753 AGCAGAAAAGGCTGAACACAGATAAATCAGCAGCAGAGAGGCGCAGTGTCTGGCGAA 812
Db 121 AGCAGAAAAGGCTGAACACAGATAAATCAGCAGCAGAGAGGCGCAGTGTCTGGCGAA 180
QY 813 GGCCAAAGGCTAAAGCTGAAGCTATTGGAATCTGGTGTGAGTCTGACACACATAATGG 872
Db 181 GGCCAAAGGCTAAAGCTGAAGCTATTGGAATCTGGTGTGAGTCTGACACACATAATGG 240
QY 873 AGATGACAGAGCTTCACTGACGTGTGCGCAGCAGATGTGTCAGCGCGTTCTCCAAACTGGC 932
Db 241 AGATGACAGAGCTTCACTGACGTGTGCGCAGCAGATGTGTCAGCGCGTTCTCCAAACTGGC 300
QY 933 CAAGGACTCCAACTATCTTACTGTGCGCTTCCAAACCTGGCGGATGTCCACAGCATGGTGGC 992
Db 301 CAAGGACTCCAACTATCTTACTGTGCGCTTCCAAACCTGGCGGATGTCCACAGCATGGTGGC 360
QY 993 TCAGGCCATGGGTGTATATGAGAGCCCTACCAAGAGCCCGCAGTGGCCAGGACTCCAGACTC 1052
Db 361 TCAGGCCATGGGTGTATATGAGAGCCCTACCAAGAGCCCGCAGTGGCCAGGACTCCAGACTC 420
QY 1053 ACTCTCCAGTGGGAGCAGCAGAGATCTCCAGGCT 1086
Db 421 ACTCTCCAGTGGGAGCAGCAGAGATCTCCAGGCT 454
RESULT 10
AX193464
LOCUS AX193464 518 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 1031 from Patent WO0149716.
ACCESSION AX193464
VERSION AX193464.1 GI:15211415
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Meagher, M.J., Xu, J. and King, G.E.
TITLE Compositions and methods for the therapy and diagnosis of colon cancer
JOURNAL Patent: WO 0173027-A 1428 04-OCT-2001;
CORIXA CORPORATION (US)
FEATURES
source Location/Qualifiers
1..518
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 123 a 136 c 149 g 109 t 1 others
ORIGIN
Query Match 30.7%; Score 365; DB 6; Length 518;
Best Local Similarity 99.4%; Pred. No. 1.5e-208;
Matches 515; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 200 CTGGGTGTGGAGCAATGGCCGATTCACCGGATCTCGAGCTGTGGAGCTGTTGAACATCC 259
Db 1 CTGGGTGTGGAGCAATGGCCGATTCACCGGATCTCGAGCTGTGGAGCTGTTGAACATCC 60
QY 260 TCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAAATGTGCATCAACG 319
Db 61 TCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAAATGTGCATCAACG 120
QY 320 TGCCTGAGCAGTGGGTGTGACTCTCGACATGTAACTCTGCNAATCGATGGAGTCTTT 379
Db 121 TGCCTGAGCAGTGGGTGTGACTCTCGACATGTAACTCTGCNAATCGATGGAGTCTTT 180
QY 380 ACTCGGCATCATGGACCTTACAAGCAAGCTACCGTGTGGAGGACCTGAGTATGCGG 439
Db 181 ACTCGGCATCATGGACCTTACAAGCAAGCTACCGTGTGGAGGACCTGAGTATGCGG 240
QY 440 TCACCCAGCTAGTCTCAACCAACCATGAGATCAGAGCTCGGCAAACTCTCTNTGAGCAAA 499
Db 241 TCACCCAGCTAGTCTCAACCAACCATGAGATCAGAGCTCGGCAAACTCTCTCTGAGCAAA 300
QY 500 TCTTCGGGAACGGGAGTCCCTGAATGCCAGCATTTGTGATGCCATCAACCAAGCTGCTG 559
Db 301 TCTTCGGGAACGGGAGTCCCTGAATGCCAGCATTTGTGATGCCATCAACCAAGCTGCTG 360
QY 560 ACTGCTGGGGTATCCCTGCTTCGTTATGAGATCAAGATATCCATGTGCCACCCCGG 619
Db 361 ACTGCTGGGGTATCCCTGCTTCGTTATGAGATCAAGATATCCATGTGCCACCCCGG 420
QY 620 TGAAGAGTCTATGACAGATGCAGGTGGAGCGAGCGCGGAAACGGCCACAGTTCTAG 679
Db 421 TGAAGAGTCTATGACAGATGCAGGTGGAGCGAGCGCGGAAACGGCCACAGTTCTAG 480
QY 680 AGTCTGAGGGGACCGAGAGTGGCCATCAATGTGGCA 717
Db 481 AGTCTGAGGGGACCGAGAGTGGCCATCAATGTGGCA 518
RESULT 11
AX261777
LOCUS AX261777 691 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 1428 from Patent WO0173027.
ACCESSION AX261777
VERSION AX261777.1 GI:16510744
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Meagher, M.J., Xu, J. and King, G.E.
TITLE Compositions and methods for the therapy and diagnosis of colon cancer
JOURNAL Patent: WO 0173027-A 1428 04-OCT-2001;

```

CORIXA CORPORATION (US)
Location/Qualifiers
1. .691
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 183 a 168 c 184 g 140 t 16 others
ORIGIN

Query Match 27.9%; Score 331; DB 6; Length 691;
Best Local Similarity 99.7%; Pred. No. 5.5e-188;
Matches 381; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 200 CCTGGTGGTGGAGCGAATGGCCGATTCACCGGATCCTGGAGCCTGGTTTGAACATCC 259
Db 1 CCTGGTGGTGGAGCGAATGGCCGATTCACCGGATCCTGGAGCCTGGTTTGAACATCC 60

Qy 260 TCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAATTTGCATCAACG 319
Db 61 TCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAATTTGCATCAACG 120

Qy 320 TGCCTGAGCAGTCGGCTGTGACTCTCGACAATGTAACTCTCAAAATCGATGGAGTCCTTT 379
Db 121 TGCCTGAGCAGTCGGCTGTGACTCTCGACAATGTAACTCTCAAAATCGATGGAGTCCTTT 180

Qy 380 ACCTGGCATCATGACCCCTTACAGGCAAGCTACGGTGTGGAGGACCCCTGAGTATGCCG 439
Db 181 ACCTGGCATCATGACCCCTTACAGGCAAGCTACGGTGTGGAGGACCCCTGAGTATGCCG 240

Qy 440 TCACCCAGCTAGCTCAACACACCATGAGATCAGAGCTCGGCAAACTCTCTTGGACAAG 499
Db 241 TCACCCAGCTAGCTCAACACACCATGAGATCAGAGCTCGGCAAACTCTCTTGGACAAG 300

Qy 500 TCTTCCGGGAACGGAGTCCCTGAATGCCAGATTTGGATGCCATCAACCAAGCTGCTG 559
Db 301 TCTTCCGGGAACGGAGTCCCTGAATGCCAGATTTGGATGCCATCAACCAAGCTGCTG 360

Qy 560 ACTGCTGGGGTATCCGCTGCCT 581
Db 361 ACTGCTGGGGTATCCGCTGCCT 382

RESULT 12
AX302706/c
LOCUS AX302706 301 bp DNA linear PAT 30-NOV-2001
DEFINITION Sequence 26 from Patent WO0179286.
ACCESSION AX302706
VERSION AX302706.1 GI:17383208
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Jiang,Y., Dillon,D.C., Mitcham,J.L., Xu,J., Harlocker,S.L. and
Repler,W.F.
TITLE Compositions and methods for the therapy and diagnosis of breast
cancer
JOURNAL Patent: WO 0179286-A 26 25-OCT-2001;
CORIXA CORPORATION (US)
Location/Qualifiers
1. .301
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 48 a 96 c 72 g 85 t
ORIGIN

Query Match 21.0%; Score 250; DB 6; Length 301;
Best Local Similarity 99.7%; Pred. No. 5.6e-139;
Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 626 AGTCTATGAGATGAGGTGGAGGAGCGCGCGAAGGCCACAGCTTCTAGAGTCTG 685
Db 626 AGTCTATGAGATGAGGTGGAGGAGCGCGCGAAGGCCACAGCTTCTAGAGTCTG 685

CORIXA CORPORATION (US)
Location/Qualifiers
1. .691
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 183 a 168 c 184 g 140 t 16 others
ORIGIN

Query Match 27.9%; Score 331; DB 6; Length 691;
Best Local Similarity 99.7%; Pred. No. 5.5e-188;
Matches 381; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 200 CCTGGTGGTGGAGCGAATGGCCGATTCACCGGATCCTGGAGCCTGGTTTGAACATCC 259
Db 1 CCTGGTGGTGGAGCGAATGGCCGATTCACCGGATCCTGGAGCCTGGTTTGAACATCC 60

Qy 260 TCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAATTTGCATCAACG 319
Db 61 TCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAATTTGCATCAACG 120

Qy 320 TGCCTGAGCAGTCGGCTGTGACTCTCGACAATGTAACTCTCAAAATCGATGGAGTCCTTT 379
Db 121 TGCCTGAGCAGTCGGCTGTGACTCTCGACAATGTAACTCTCAAAATCGATGGAGTCCTTT 180

Qy 380 ACCTGGCATCATGACCCCTTACAGGCAAGCTACGGTGTGGAGGACCCCTGAGTATGCCG 439
Db 181 ACCTGGCATCATGACCCCTTACAGGCAAGCTACGGTGTGGAGGACCCCTGAGTATGCCG 240

Qy 440 TCACCCAGCTAGCTCAACACACCATGAGATCAGAGCTCGGCAAACTCTCTTGGACAAG 499
Db 241 TCACCCAGCTAGCTCAACACACCATGAGATCAGAGCTCGGCAAACTCTCTTGGACAAG 300

Qy 500 TCTTCCGGGAACGGAGTCCCTGAATGCCAGATTTGGATGCCATCAACCAAGCTGCTG 559
Db 301 TCTTCCGGGAACGGAGTCCCTGAATGCCAGATTTGGATGCCATCAACCAAGCTGCTG 360

Qy 560 ACTGCTGGGGTATCCGCTGCCT 581
Db 361 ACTGCTGGGGTATCCGCTGCCT 382

RESULT 13
AX260945
LOCUS AX260945 740 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 596 from Patent WO0173027.
ACCESSION AX260945
VERSION AX260945.1 GI:16509912
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Meagher,M.J., Xu,J. and King,G.E.
TITLE Compositions and methods for therapy and diagnosis of colon cancer
JOURNAL Patent: WO 0173027-A 596 04-OCT-2001;
CORIXA CORPORATION (US)
Location/Qualifiers
1. .740
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 189 a 181 c 169 g 159 t 42 others
ORIGIN

Query Match 14.2%; Score 169; DB 6; Length 740;
Best Local Similarity 99.5%; Pred. No. 5.8e-90;
Matches 219; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 247 GGTTTGAACATCCTCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAA 306
Db 120 GGTTTGAACATCCTCATCCCTGTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAA 179

Qy 307 ATTGTATCAACGTCCTGAGCAGTCGGCTGTGACTCTCGACAATGTAACTCTGCAATC 366
Db 180 ATTGTATCAACGTCCTGAGCAGTCGGCTGTGACTCTCGACAATGTAACTCTGCAATC 239

Qy 367 GATGGAGTCCTTTACCTGGGCATCATGGACCTTACAGGCAAGCTACGGTGTGGAGGAC 426
Db 240 GATGGAGTCCTTTACCTGGGCATCATGGACCTTACAGGCAAGCTACGGTGTGGAGGAC 299

Qy 427 CCTGAGTATGCCGTCAACCCAGCTAGCTCAACCAACCATGA 466
Db 300 CCTGAGTATGCCGTCAACCCAGCTAGCTCAACCAACCATGA 339

RESULT 14
AC026501/c
LOCUS AC026501 55406 bp DNA linear HTG 22-MAR-2000
DEFINITION Homo sapiens chromosome 1 clone RP11-266M12 map 1, LOW-PASS
```

```

SEQUENCE SAMPLING.
AC026501
VERSION AC026501.1 GI:7283191
KEYWORDS HTG: HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 55406)
JOURNAL Birren,B., Linton,L., Nusbaum,C. and Lander,E.
REFERENCE Homo sapiens chromosome 1, clone RP11-266M12
AUTHORS Unpublished
2 (bases 1 to 55406)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campolavio,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Glnde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczyk,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McCurk,A., McKernan,K., McPheeters,R.,
Melidrim,J., Meneus,D., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
ROY,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Triglio,J.,
Vassiliou,H., Viel,R., Vo.A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center -----
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information -----
Center project name: L6433
Center clone name: 266_M_12
-----
* NOTE: This record contains 70 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1
* 677 776: contig of 676 bp in length
* 777 776: gap of 100 bp
* 1478 1478: contig of 702 bp in length
* 1578 1578: gap of 100 bp
* 2267 2267: contig of 689 bp in length
* 2367 2367: gap of 100 bp
* 3052 3052: contig of 685 bp in length
* 3152 3152: gap of 100 bp
* 3870 3870: contig of 718 bp in length
* 3970 3970: gap of 100 bp
* 4671 4671: contig of 701 bp in length
* 4771 4771: gap of 100 bp
* 5475 5475: contig of 704 bp in length
*
* 5476 5575: gap of 100 bp
* 5576 6263: contig of 688 bp in length
* 6264 6363: gap of 100 bp
* 6364 7060: contig of 697 bp in length
* 7061 7160: gap of 100 bp
* 7161 7869: contig of 709 bp in length
* 7870 7969: gap of 100 bp
* 7970 8683: contig of 714 bp in length
* 8684 8783: gap of 100 bp
* 8784 9466: contig of 683 bp in length
* 9467 9566: gap of 100 bp
* 9567 10263: contig of 697 bp in length
* 10264 10363: gap of 100 bp
* 10364 11053: contig of 690 bp in length
* 11054 11153: gap of 100 bp
* 11154 11848: contig of 695 bp in length
* 11849 11948: gap of 100 bp
* 11949 12639: contig of 691 bp in length
* 12640 12739: gap of 100 bp
* 12740 13444: contig of 705 bp in length
* 13445 13544: gap of 100 bp
* 13545 14237: contig of 693 bp in length
* 14238 14337: gap of 100 bp
* 14338 15047: contig of 710 bp in length
* 15048 15147: gap of 100 bp
* 15148 15720: contig of 573 bp in length
* 15721 15820: gap of 100 bp
* 15821 16512: contig of 692 bp in length
* 16513 16612: gap of 100 bp
* 16613 17316: contig of 704 bp in length
* 17317 17416: gap of 100 bp
* 17417 18118: contig of 702 bp in length
* 18119 18218: gap of 100 bp
* 18219 18904: contig of 686 bp in length
* 18905 19004: gap of 100 bp
* 19005 19696: contig of 692 bp in length
* 19697 19796: gap of 100 bp
* 19797 20486: contig of 690 bp in length
* 20487 20586: gap of 100 bp
* 20587 21293: contig of 707 bp in length
* 21294 21393: gap of 100 bp
* 21394 22098: contig of 705 bp in length
* 22099 22198: gap of 100 bp
* 22199 22875: contig of 677 bp in length
* 22876 22975: gap of 100 bp
* 22976 23682: contig of 707 bp in length
* 23683 23782: gap of 100 bp
* 23783 24474: contig of 692 bp in length
* 24475 24574: gap of 100 bp
* 24575 25280: contig of 706 bp in length
* 25281 25380: gap of 100 bp
* 25381 26083: contig of 703 bp in length
* 26084 26183: gap of 100 bp
* 26184 26874: contig of 691 bp in length
* 26875 26974: gap of 100 bp
* 26975 27679: contig of 705 bp in length
* 27680 27779: gap of 100 bp
* 27780 28457: contig of 678 bp in length
* 28453 28557: gap of 100 bp
* 28558 29255: contig of 698 bp in length
* 29256 29355: gap of 100 bp
* 29356 30053: contig of 698 bp in length
* 30054 30153: gap of 100 bp
* 30154 30847: contig of 694 bp in length
* 30848 30947: gap of 100 bp
* 30948 31656: contig of 709 bp in length
* 31657 31756: gap of 100 bp
* 31757 32459: contig of 703 bp in length
* 32460 32559: gap of 100 bp
* 32560 33270: contig of 711 bp in length
* 33271 33370: gap of 100 bp
* 33371 34072: contig of 702 bp in length
* 34073 34172: gap of 100 bp

```


Job time: 7130 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 16:14:48 ; Search time 209.97 Seconds
(without alignments)
9714.218 Million cell updates/sec

Title: US-09-898-216-2

Perfect score: 1188

Sequence: 1 GGCTTCTGGAGCNACCGCT.....GGAAGCAGATTTTCCTGATT 1188

Scoring table: OLIGO NUC

Learning curve: 0.000000
Gapop 60.0 , Gapext 60.0

[illegible]

Searched: 1736436 seqs, 858457221 residues

1000

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

```

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_Q32802.*
1: /SIDSL/cgcdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSL/cgcdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSL/cgcdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSL/cgcdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSL/cgcdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSL/cgcdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSL/cgcdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSL/cgcdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSL/cgcdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSL/cgcdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSL/cgcdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSL/cgcdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSL/cgcdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSL/cgcdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSL/cgcdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSL/cgcdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSL/cgcdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSL/cgcdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSL/cgcdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSL/cgcdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSL/cgcdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSL/cgcdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSL/cgcdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSL/cgcdata/geneseq/geneseqn-emb1/NA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	1188	100.0	1188	19	AAV28867	Human integral mem	
2	857	72.1	1337	20	AAAX04345	Human secreted pro	
C 3	857	72.1	1398	22	ABA09225	Human SLP-2 homolo	
C 4	857	72.1	1398	22	AAK52683	Human polynucleoti	
5	857	72.1	9098	22	AAK54953	cDNA encoding nove	
6	855	72.0	1322	20	AAAX04402	Human secreted pro	
7	806	67.8	1244	22	AAH13961	Human cDNA sequenc	
8	793	66.8	1416	22	AAK516599	Human polynucleoti	
9	365	30.7	518	22	AAI29477	Colon tumour relat	

C	10	335	28.2	566	22	AAK93925	Human cDNA clone
	11	331	27.5	691	23	AAK58752	cDNA #1428 encoding
	12	285	24.0	2064	22	AAK94818	Human full-length
C	13	253	21.3	313	22	AAK23495	Human breast cancer
C	14	250	21.0	301	22	AAK46972	Human breast cancer
C	15	250	21.0	301	22	AAK17542	Human breast cancer
	16	217	18.3	761	22	AAK24446	Human breast cancer
C	17	192	16.2	421	22	AAK18944	Human polynucleoti
C	18	176	14.8	565	22	AAK92831	Human cDNA 3'-end
C	19	176	14.8	565	22	AAK94037	Human cDNA clone r
	20	171	14.4	567	22	AAK93152	Human cDNA clone r
	21	171	14.4	567	22	AAH05115	Human cDNA clone (
	22	169	14.2	740	23	AAK57920	cDNA #596 encoding
C	23	124	10.4	417	22	AAK15602	Human breast cancer
C	24	122	10.3	607	22	AAK14633	Human breast cancer
	25	84	7.1	207	23	AAK58068	cDNA #744 encoding
	26	82	6.9	190	21	AAK17814	Human secreted pro
	27	74	6.2	253	15	AAQ16727	Human genome fragm
	28	67	5.6	202	21	AAK242655	Human 5' EST isola
	29	66	5.6	1682	22	AAK45141	cDNA encoding nove
	30	40	3.4	163	16	AAK19309	Human gene signatu
	31	40	3.4	747	20	AAK215045	Human gene express
	32	37	3.1	300	20	AAK13550	Human gene express
	33	37	3.1	300	20	AAK98374	Human cancer cell
	34	21	1.8	28598	17	AAK06769	Sorangium cellulos
	35	21	1.8	28958	18	AAK89956	Sorangium cellulos
	36	21	1.8	28958	21	AAK75299	DNA sequence of So
	37	21	1.8	49377	19	AAK05287	The soraphen biosy
C	38	20	1.7	93	22	ABA76044	Human foetal liver
C	39	20	1.7	93	22	ABA40604	Probe #19070 for g
C	40	20	1.7	93	22	AAK24720	Human brain expres
C	41	20	1.7	93	22	AAK50720	Human bone marrow
C	42	20	1.7	93	22	AAK127734	Probe #17667 for g
C	43	20	1.7	93	22	AAK156701	Probe #25387 used
C	44	20	1.7	548	22	ABA63763	Human foetal liver
C	45	20	1.7	548	22	ABA30955	Probe #9421 for ge

ALIGNMENTS

RESULT 1

AAV28867

ID AAV28867 standard; DNA; 1188 BP.

AAV28867;

03-AUG-1998 (first entry)

Human integral membrane protein encoding DNA.

Human; integral membrane protein; IMP; cancer; anaemia; prostate; breast; pancreatic; tumour; ion transport; ss.

Homo sapiens.

Key

Location/Qualifiers

64..1134

/*tag= a

/product= "integral membrane protein"

/transl_except= (pos:490..492,aa:Xaa)

/transl_except= (pos:1099..1101,aa:Xaa)

/note= "Xaa= unknown"

US5763589-A.

09-JUN-1998.

09-JAN-1997; 97US-0781562.

09-JAN-1997; 97US-0781562.

(INCV-) INCVTE PHARM INC

PR 13-JUN-1997; 97US-0050901.
 PR 13-JUN-1997; 97US-0052989.
 PR 08-JUL-1997; 97US-0051919.
 PR 18-AUG-1997; 97US-0055984.
 PR 12-SEP-1997; 97US-0058665.
 PR 12-SEP-1997; 97US-0058668.
 PR 12-SEP-1997; 97US-0058669.
 PR 12-SEP-1997; 97US-0058750.
 PR 12-SEP-1997; 97US-0058971.
 PR 12-SEP-1997; 97US-0058972.
 PR 12-SEP-1997; 97US-0058975.
 PR 02-OCT-1997; 97US-0060834.
 PR 02-OCT-1997; 97US-0060841.
 PR 02-OCT-1997; 97US-0060844.
 PR 02-OCT-1997; 97US-0060865.
 PR 02-OCT-1997; 97US-0061059.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Brewer LA, Ebner R, Ferrie AM, Feng P, Greene JM, Lafleur DW;
 PI Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM, Shi Y, Young P;
 PI Yu GL;
 XX
 DR WPI: 1999-080881/07.
 DR P-PSDB; AAW78160.
 XX
 PT New isolated human genes and the secreted polypeptides they encode -
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders
 XX
 PS Claim 1; Page 201-202; 380pp; English.
 XX
 CC This sequence represents a nucleic acid molecule which encodes a
 CC secreted human protein. The gene number, and the clone it is derived
 CC from, are detailed in the descriptor line. The gene can be used to
 CC generate fusion proteins by linking to the gene to a human immunoglobulin
 CC Fc portion (e.g. AAX04302) for increasing the stability of the fused
 CC protein as compared to the human protein only.
 CC The invention relates to 86 novel genes and their fragments (nucleic
 CC acid sequences: AAX04311-X04410; amino acid sequences AAW78126-W78225)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 86
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AAX04311 for described uses).
 XX
 SQ Sequence 1337 BP; 338 A; 346 C; 377 G; 273 T; 3 other;

Query Match 72.1%; Score 857; DB 20; Length 1337;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 140 GCGGCGCTCTCTGTGATTGCCCGCAACACCGTGTACTGTTCGGCCGACGAGAGG 199
 DB |||||||
 DB 152 GCGGCGCTCTCTGTGATTGCCCGCAACACCGTGTACTGTTCGGCCGACGAGAGG 211
 QY 200 CTGGGTGTGGAGCAATGGCCGATTCACCGGATCTCGAGCTGCTGTTTCAACATCC 259
 DB |||||||
 DB 212 CTTG99GT99GAGCGAAAGG99CGGATCCCGGATCTGAGCTG99GTTTGAACATCC 271
 QY 260 TCATCCTCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAAATGTTCATCAACG 319
 DB |||||||
 DB 272 TCATCCTCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAAATGTTCATCAACG 331
 QY 320 TGCTGAGCAGTGGCGTGTGACTCTCGAATGTAACTGTCAAAATCGATGAGTCCCTTT 379
 DB |||||||
 DB 332 TGCTGAGCAGTGGCGTGTGACTCTCGAATGTAACTGTCAAAATCGATGAGTCCCTTT 391
 QY 380 ACCTGGCATATGACCCCTTACAGGCAAGCTACGGTGTGGAGGACCTGAGTATGCCG 439
 DB |||||||

Db 392 acctgogcatcatgacccctacaaggcgaagtacgtgtgtgaggaccctgagatgctg 451
 QY 440 TCACCCAGCTAGTCAACAAACCATGAGATCAGAGCTCGGCAAACTCTCTNTGGACAAAG 499
 Db |||||||
 Db 452 tcaccagctagtctcaaacacacatgagatcagagctcgcaaacctctctcttgacaaa 511
 QY 500 TCTTCGGGAAAGGGAGTCCCTGAATGCGAGCATTTGGATGCCATCAACCAAGCTGCTG 559
 Db |||||||
 Db 512 tcttcgggaaacgggagtgccctgaatgcccagcatttggtatgcatcaaccagagctgtg 571
 QY 560 ACTGCTGGGTTATCCGCTCCCTTCGTTATGATCAAGGATATCCATGTGCCACCCCGGG 619
 Db |||||||
 Db 572 actgtg99gtatccgtcccttatgadatcaagatatccatgtgcccaccccg99 631
 QY 620 TGAAGAGTCTATGAGATGCGAGTGCAGGCGAGCGGGGAAACGGCCACACAGTCTCTAG 679
 Db |||||||
 Db 632 tgaagagctctatgcatgagtgagtgagcagagcg99gaaacgggcccacagttctag 691
 QY 680 AGTCTAGGGGCCCGGAGAGTGGCCATCAATGTGCGAGGAAGGAAGAACAGGCCGAGA 739
 Db |||||||
 Db 692 agtctgaggggacccgagagtgcccatcaatgtgagaaagggaagaaacagggcccaga 751
 QY 740 TCCTGGCTCCGGAAGCAGAAAGGCTGAACAGATAAATCAGGCGAGGAGGCCAGTG 799
 Db |||||||
 Db 752 tcttgccctccgaagcagaaaggctgaacagataaatcaggcagcagagagggccagt 811
 QY 800 CAGTTCCTGGCGAGGCCAAGGCTAAAGCTGAAGCTATTCGAATCTCGCTGCGAGCTCTGA 859
 Db |||||||
 Db 812 cagttctg99gaaaggccaaaggctaaagctgaagctattcgaatcctggctgagctctga 871
 QY 860 CACAACATAATGGAGATGCGAGCTTCACTGACTGTGGCCGAGCAGTATGTTCAGGCGCT 919
 Db |||||||
 Db 872 cacaacataatggagatgagcagcagcttctactgtgtg99cagcagatgtgcagcgct 931
 QY 920 TCTCCAACTGGCCAAAGGACTCCACACATATCTACTGCTCCCAACCTTGGCGATGTCA 979
 Db |||||||
 Db 932 tctccaaactggcaggagctccactctactgcccctccacccctgg99cagtgca 991
 QY 980 CCAGCATGTGTGCTCAGGCCATGGGTGTATATGAGCCCTCACCAGCCCGAGTCCAG 1039
 Db |||||||
 Db 992 ccagcagtggtgctcagcagcagtggtgtatatgagccctcaccacagcccagtcgag 1051
 QY 1040 GGACTCCAGCTCACTCTCCAGTGGGAGGAGCAGAGATGTCAGGGGTACAGATGCAATG 1098
 Db |||||||
 Db 1052 ggactccagactcactctccagtg99gagcagcagatgtccagggtacagatgcaagt 1110
 RESULT 3
 ABA09225/c
 ID ABA09225 standard; cDNA; 1398 BP.
 XX
 AC ABA09225;
 XX
 DT 11-JAN-2002 (first entry)
 XX
 DE Human SLP-2 homologue-encoding cDNA, SEQ ID NO:1001.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibitor; chemotaxis; chemokines; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnerary; antiulcer; ss.
 OS Homo sapiens.
 XX

PN WO200157188-A2.
 XX 09-AUG-2001.
 XX 05-FEB-2001; 2001WO-US03800.
 XX 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT;
 PI WPI; 2001-457740/49.
 DR P-PSDB; ABB11981.
 XX Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 XX e.g. arthritis and cancer -
 XX
 PS Claim 1; Page 851; 1963pp; English.
 XX
 CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a cDNA encoding a
 CC novel human polypeptide of the invention.
 XX
 XX Sequence 1398 BP; 307 A; 408 C; 359 G; 324 T; 0 other;

Query Match 72.1%; Score 857; DB 22; Length 1398;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 140 GCCGGCCCTCTGGATTGCCCGAACCACCGTGGTACTGTCGCCGAGCAGGAGG 199
 Db 1254 GCCGGCCCTCTGGATTGCCCGAACCACCGTGGTACTGTCGCCGAGCAGGAGG 1195
 Qy 200 CTGGGTGGTGGAGCGAATGGCGCATTCACCGGATCCTGGAGCGCTGGTTGAACATCC 259
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 1194 CCTGGTGTGGAGCGAATGGCGCATTCACCGGATCCTGGAGCGTGTGTTGAACATCC 1135
 Qy 260 TCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGAAATTTGTCATCAACG 319
 Db 1134 TCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGAAATTTGTCATCAACG 1075
 Qy 320 TGCCTGAGCAGTCGGCTGTGACTCTCGACAATGTAACTCTGCAAAATCGATGAGTCCCTTT 379
 Db 1074 TGCCTGAGCAGTCGGCTGTGACTCTCGACAATGTAACTCTGCAAAATCGATGAGTCCCTTT 1015
 Qy 380 ACCTCGGCATCATGACCCCTTACAAGGCAAGCTACGGTGTGGAGGACCTCTCTNTGGACAAG 439
 Db 1014 ACCTCGGCATCATGACCCCTTACAAGGCAAGCTACGGTGTGGAGGACCTCTCTNTGGACAAG 955
 Qy 440 TCACCCAGCTAGCTCAAAACAACCATGAGATCAGAGCTCGGCAAACTCTCTNTGGACAAG 499
 Db 954 TCACCCAGCTAGCTCAAAACAACCATGAGATCAGAGCTCGGCAAACTCTCTNTGGACAAG 895
 Qy 500 TCTTCCGGGAACGGGAGTCCCTGTAATGCCAGCATTTGTGGATGCCATCAACCAAGCTGCTG 559
 Db 894 TCTTCCGGGAACGGGAGTCCCTGTAATGCCAGCATTTGTGGATGCCATCAACCAAGCTGCTG 835
 Qy 560 ACTGTGGGGTATCCGCTGCGCTTATGAGATCAAGATATCCATGTGCCACCCCGGG 619
 Db 834 ACTGTGGGGTATCCGCTGCGCTTATGAGATCAAGATATCCATGTGCCACCCCGGG 775
 Qy 620 TGAAGAGTCTATGATGAGTGGAGGAGCGGCGGAAAGCGGCGGAGGAGGAGGAGGAGGAGG 679
 Db 774 TGAAGAGTCTATGATGAGTGGAGGAGCGGCGGAAAGCGGCGGAGGAGGAGGAGGAGGAGG 715
 Qy 680 AGTCTGAGGGACCGGAGAGTGGCGCATCAATGTGGCAGAAGGAGGAGGAGGAGGAGGAGGAGG 739
 Db 714 AGTCTGAGGGACCGGAGAGTGGCGCATCAATGTGGCAGAAGGAGGAGGAGGAGGAGGAGGAGG 655
 Qy 740 TCCTGGCTCCGAGCAAGAAAGGCTGAACAGATTAATCAGGACGAGGAGGAGGAGGAGGAGGAGG 799
 Db 654 TCCTGGCTCCGAGCAAGAAAGGCTGAACAGATTAATCAGGACGAGGAGGAGGAGGAGGAGGAGG 595
 Qy 800 CAGTTCTCGGAGCGCCAGGCTAAAGCTGAAGCTATTCGAATCTGGCTGCAGCTCTGA 859
 Db 594 CAGTTCTCGGAGCGCCAGGCTAAAGCTGAAGCTATTCGAATCTGGCTGCAGCTCTGA 535
 Qy 860 CACAACATAATGAGATCGACAGCTTCACTGACTGTGGCCGAGCAGTATGTGACGGCT 919
 Db 534 CACAACATAATGAGATCGACAGCTTCACTGACTGTGGCCGAGCAGTATGTGACGGCT 475
 Qy 920 TCTCCAAACTGCCAAGACATCCACACTATCTTACTGCCCTCCAAACCTGCCGATGTCA 979
 Db 474 TCTCCAAACTGCCAAGACATCCACACTATCTTACTGCCCTCCAAACCTGCCGATGTCA 415
 Qy 980 CCAGCATGGTGGCTCAGGCCATGGGTGTATATGGAGCCCTCACCAGGAGGAGGAGGAGGAGGAGG 1039
 Db 414 CCAGCATGGTGGCTCAGGCCATGGGTGTATATGGAGCCCTCACCAGGAGGAGGAGGAGGAGGAGG 355
 Qy 1040 GGACTCCAGATCACTCTCCAGTGGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGT 1098
 Db 354 GGACTCCAGATCACTCTCCAGTGGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGT 296

RESULT 4
 AAK52683/c
 ID AAK52683 standard; cDNA; 1398 BP.
 XX
 AC AAK52683;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 2212.
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.

XX Homo sapiens.
XX WO200157190-A2.
XX
XX
XX PD 09-AUG-2001.
XX PF
XX 05-FEB-2001; 2001WO-US04098.
XX
XX 03-FEB-2000; 2000US-0496914.
XX PR 27-APR-2000; 2000US-0560875.
XX PR 20-JUN-2000; 2000US-0598075.
XX PR 19-JUL-2000; 2000US-0620325.
XX PR 01-SEP-2000; 2000US-0654936.
XX PR 15-SEP-2000; 2000US-0663561.
XX PR 20-OCT-2000; 2000US-0693325.
XX PR 30-NOV-2000; 2000US-0728422.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX WPI: 2001-476283/51.
XX DR P-P5DB; AAM79550.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
XX Claim 1: Page 4572; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM8020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
XX Sequence 1398 BP; 307 A; 408 C; 359 G; 324 T; 0 other;
SQ

Query Match 72.1%; Score 857; DB 22; Length 1398;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 140 GCGGCGCTCTGTGGATTGCCCGGAAACACCGTGTACTGTTCGTGCGCAGCAGGAGG 199
DB 1254 GCGGCGCTCTCTGTGGATTGCCCGGAAACACCGTGTACTGTTCGTGCGCAGCAGGAGG 1195
QY 200 CTGTGGTGTGGAGCGAATGGCGGATTCACCGGATCTCGAGCCTGTGTTGAACATCC 259
DB 1194 CTGTGGTGTGGAGCGAATGGCGGATTCACCGGATCTCGAGCCTGTGTTGAACATCC 1135
QY 260 TCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAAATGTGCATCAACG 319
DB 1134 TCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAAATGTGCATCAACG 1075
QY 320 TGCCTGAGCAGTCGGTGTGACTCTCGACAATGTAACTCTGCAATTCGATGGAGTCCTTT 379
DB 1074 TGCCTGAGCAGTCGGTGTGACTCTCGACAATGTAACTCTGCAATTCGATGGAGTCCTTT 1015
QY 380 ACCTGGGCATCATGACCCCTTACAGGCAAGCTACGGTGTGGAGGACCTGAGTATGCCG 439
DB 1014 ACCTGGGCATCATGACCCCTTACAGGCAAGCTACGGTGTGGAGGACCTGAGTATGCCG 955

QY 440 TCACCCAGCTAGCTCAACAACCATGAGATCAGAGCTCGGCAAACTCTCTTNTGGACAAAG 499
DB 954 TCACCCAGCTAGCTCAACAACCATGAGATCAGAGCTCGGCAAACTCTCTCTGGACAAAG 895
QY 500 TCTTCCGGGAAAGGGAGTCCCTGAATGCCAGCATTTGTGATGCCATCAACCAAGCTGCTG 559
DB 894 TCTTCCGGGAAAGGGAGTCCCTGAATGCCAGCATTTGTGATGCCATCAACCAAGCTGCTG 835
QY 560 ACTGTGCGGTATCCGCTGCCTTCGCTTATGAGATCAAGATATCCATGTCACCCCGGG 619
DB 834 ACTGTGCGGTATCCGCTGCCTTCGCTTATGAGATCAAGATATCCATGTCACCCCGGG 775
QY 620 TGAAGAGTCTATGCAGATGCAGGTGGAGCGAGCGGGGAAACGGGCCACATTTCTAG 679
DB 774 TGAAGAGTCTATGCAGATGCAGGTGGAGCGAGCGGGGAAACGGGCCACATTTCTAG 715
QY 680 AGTCTGAGGGGACCCGAGAGTCCGCTCATCAATGTGCGAGAGGAAGAAGACAGCCCGA 739
DB 714 AGTCTGAGGGGACCCGAGAGTCCGCTCATCAATGTGCGAGAGGAAGAAGACAGCCCGA 655
QY 740 TCGTGGCTCCCAAGCAGAAAGGCTGAACAGATAAATCAGGCAGCAGGAGGCGCAGTG 799
DB 654 TCGTGGCTCCCAAGCAGAAAGGCTGAACAGATAAATCAGGCAGCAGGAGGCGCAGTG 595
QY 800 CAGTTCTGCGAAGGCCAAAGGCTAAAGCTGAAGCTATTCGAATCCTGGCTGCAAGTCTGA 859
DB 594 CAGTTCTGCGAAGGCCAAAGGCTAAAGCTGAAGCTATTCGAATCCTGGCTGCAAGTCTGA 535
QY 860 CACAACATATGAGATGCAGCAGCTTCACTGACTGTGCGCGAGCAGTATGTGAGCGCT 919
DB 534 CACAACATATGAGATGCAGCAGCTTCACTGACTGTGCGCGAGCAGTATGTGAGCGCT 475
QY 920 TCTCCAACTGGCCAAAGGACTCCCAACACTATCTACTGCGCTCCAACTCCCTGGCGATGTCA 979
DB 474 TCTCCAACTGGCCAAAGGACTCCCAACACTATCTACTGCGCTCCAACTCCCTGGCGATGTCA 415
QY 980 CCAGCATGTGGCTCAGGCGCATGGGTGTATATGAGCGCTTCACCAAGCCCGCAGTGCCAG 1039
DB 414 CCAGCATGTGGCTCAGGCGCATGGGTGTATATGAGCGCTTCACCAAGCCCGCAGTGCCAG 355
QY 1040 GGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGTCCAGGCTACAGATGCAAGT 1098
DB 354 GGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGTCCAGGCTACAGATGCAAGT 296

RESULT 5
AAS44953
ID AAS44953 standard; cdna; 9098 BP.
XX
XX AAS44953;
XX AC
XX DT 18-DEC-2001 (first entry)
XX DE
XX CDNA encoding novel human secretory protein, Seq ID No 34.
XX
XX Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
KW gut protection; lung; liver fibrosis; immune deficiency; infection;
KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
KW fertility; analgesic; pain; antigen; ss.
XX
XX Homo sapiens.
XX
XX WO200166689-A2.
XX
XX 13-SEP-2001.
XX
XX 05-MAR-2001; 2001WO-US04942.

XX Homo sapiens.
OS WO9856804-A1.
PN 17-DEC-1998.
PD 11-JUN-1998;
PF 98WO-US12125.
XX 02-OCT-1997;
XX 13-JUN-1997;
PR 97US-0061060.
PR 97US-0049547.
PR 97US-0049548.
PR 97US-0049549.
PR 97US-0049550.
PR 97US-0049606.
PR 97US-0049607.
PR 97US-0049608.
PR 97US-0049609.
PR 97US-0049610.
PR 97US-0049611.
PR 97US-0050566.
PR 97US-0050901.
PR 97US-0052989.
PR 08-JUL-1997;
PR 18-AUG-1997;
PR 12-SEP-1997;
PR 12-SEP-1997;
PR 12-SEP-1997;
PR 12-SEP-1997;
PR 12-SEP-1997;
PR 12-SEP-1997;
PR 02-OCT-1997;
PR 02-OCT-1997;
PR 02-OCT-1997;
XX (HUMA-) HUMAN GENOME SCI INC.
XX Brewer LA, Ebner R, Ferrrie AM, Feng P, Greene JM, Lafleur DW;
PI Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM, Shi Y, Young P;
PI Yu GL;
XX WPI: 1999-080881/07.
DR P-PSDB; AAW78217.
XX New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
XX Claim 1: Page 257-258; 380pp; English.
PS This sequence represents a nucleic acid molecule which encodes a
XX secreted human protein. The gene number, and the clone it is derived
CC from, are detailed in the descriptor line. The gene can be used to
CC generate fusion proteins by linking to the gene to a human immunoglobulin
CC Fc portion (e.g. AAX04302) for increasing the stability of the fused
CC protein as compared to the human protein only.
CC The invention relates to 86 novel genes and their fragments (nucleic
CC acid sequences: AAX04311-X04410; amino acid sequences AAW78126-W78225)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 86
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AAX04311 for described uses).
XX Sequence 1322 BP; 336 A; 337 C; 373 G; 271 T; 5 other;
SQ

Query Match 72.0%; Score 855; DB 20; Length 1322;

Best Local Similarity 99.8%; Pred. No. 0;
Matches 955; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 142 CGCGCTCTCTGGATTGCCCGGAAACACCGCTGGTACTGTTCTGTCGCGCAGCAGGAGGCC 201
DB 139 cgcgcctcctctggattgcccgcaaacaccgctgggtactgttctcgtccgcagcaggaagcgc 198
QY 202 TGGGTGGTGGAGCGCAATGGCGCATTTCCACCGGATTCCTACCGGCTGGTGGTGAACATPCTTC 261
DB 199 tgggtggtagcgcaatggcgcatccacccgattccacccgattccctggagcctggtttgaacatcctc 258
QY 262 ATCCCTGTGTAGACCGGATTCGATATGTGAGAGTCTCAAGGAAATTTCTCATCAACGTG 321
DB 259 atccctgtgttagaccggatccgatattgacagagctctcaagaaatttctcatcaacgtg 318
QY 322 CCTGAGCAGTCTGGCTGTGACTCTCGACAATGTAACCTCTCAATTCGATCGAGTCTCTTAC 381
DB 319 cctgagcagtcggctgtgactctcgaactctgaactctcgaactctcctctctggacaaagtc 378
QY 382 CTGCGCATCATGACCCCTTACAGGCAAGCTACGGTGTGGAGACCCCTGAGTATGCCGTG 441
DB 379 ctgcgcatcatggacccttacaaggcaagctacggctgtaggagaccctgagtatgscgtc 438
QY 442 ACCCAGCTAGCTCAACACCATGAGATCAGAGTCTGGCGAAACTCTCTTTGGCAAAAGTC 501
DB 439 acccagctagctcaaacacacatgagatcagagctcggcaactctctctctggacaaagtc 498
QY 502 TTCCGGGAACGGGAGTCCCTGAATGCCAGCATTTGTGGATGCCATCAACCAAGCTGCTGAC 561
DB 499 ttccgggaacgggagtcctctgaatgccagcatctggtggatcaacaaagctgctgac 558
QY 562 TGCTGGGTATCCGCTCCCTGTTATGAGATCAAGGATATCCATGTGCCACCCCGGTG 621
DB 559 tgctgggtatccgctgctcctgttatgagatcaagatatccatgtgccaccccggtg 618
QY 622 AAGAGTCTATGAGATGAGGTGGAGGAGCGCGGAAACGGGCCACAGTCTTAGAG 681
DB 619 aaagagctctatgagatgaggtggaggcagagcggcaaacggccacagttctagag 678
QY 682 TCTGAGGCGACCCGAGAGTCCGCTCAATGTGGCAGAGAGGGAAGAAACAGGCCACAGATC 741
DB 679 tctgaggggagcccgagagctcgccca tcaatgtggcaggaagggaagaaacagggcccgatc 738
QY 742 CTGGCTCCGAGCAGAGAAAGGCTGAACAGATAAATCAGSCAGCAGGAGGCCAGTGCA 801
DB 739 ctggcctccgagcagaaaaggctgaacagataaatcaggcagcagagagggccagtgca 798
QY 802 GTTCTGGCGAAGGCCAAGGCTTAAGCTGAAGCTATTTCGAATCTGCTGCTGAGTCTTGACA 861
DB 799 gttctggcgaaggccaaggctaaagctgagctattcgaatcctggtgctgagctctgaca 858
QY 862 CAACATAATGGAGATGCAGCAGCTTCACTGTGGCGAGCAGTATGTACGCGGTTC 921
DB 859 caacataatggagatgcagcagcttcactgactgtggcgagcagtatgtcagcgagttc 918
QY 922 TCCAAACTGGCCCAAGGACTCCAACTATCTCTACTCCCTTCCAACTCCGCGATGTCACC 981
DB 919 tccaaactggcgaaggactccaaactatctactgctccctcccaacctccgctgctgacc 978
QY 982 AGCATGCTGGCTCAGGCCATGGGTGTATATGGAGCCCTCACCAGGCCCGCCAGG 1041
DB 979 agcatggtgctcagggccatgggtgtatggagccctcaccagaagccccagtgccagg 1038
QY 1042 ACTCCAGACTCTCTCTCCAGTGGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGT 1098
DB 1039 actccagactcactctccagtggggagcagcagagatgtccaggggtacagatgcaagt 1095
RESULT 7
AAH13961
ID AAH13961 standard; cDNA; 1244 BP.
XX
AC AAH13961;

XX
DT 26-JUN-2001 (first entry)
XX Human cDNA sequence SEQ ID NO:11009.
DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
KW Homo sapiens.
XX EP1074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-0116126.
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-0118776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX Claim 8; SEQ ID 11009; 2537pp + CD ROM; English.
XX The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX Sequence 1244 BP; 284 A; 327 C; 368 G; 265 T; 0 other;
SQ

Query Match 67.8%; Score 806; DB 22; Length 1244;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 956; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 140 GCCGGCTCTCTCGATGTCCTCCCAACACCGTGGTACTCTGTCGCCGAGCAGGAGG 199
Db 111 gccgcgcctctctgattgctcccgaaacaccgtggtactgttcgtccgcagcaggagg 170
Qy 200 CTGGGTGGTGGAGCGAATGGCGGATTCACCGGATCTCTGGAGCCTGGTTTGAACATCC 259
Db 171 cctgggtggtgagcgaatggcgcattccaccggatcctcgagcctggttgaaatccc 230

QY	260	TCATCCCTGTTAGACCGGATCCGATATGTCAGAGTCTCAAGGAAATTTGTATCAACG	319
DB	231	tcataccctgtttagaccggaatccgataatgtagaggtctcaaggaaatctgcatcaacg	290
QY	320	TGCTGAGCAGTCGGCTGTGACTCTCGACAATTAAGTAACTCTGAAATCGATGGAGTCTCTT	379
DB	291	tgactgagcagtcgctgtgactctcgacaataaactctgcaaatcgatggagtccttt	350
QY	380	ACCTGGCGATCATGAGACCCCTTACAAGCAAGCTACGGTGTGGAGGACCCCTGATATGCCG	439
DB	351	acctgcgcatcatggacccttaacaggcaagctacggtgtggaggaccctgagtatgcg	410
QY	440	TCACCCAGCTAGCTCAACCAACCATGAGATCAGAGTCGCGCAAACTCTCTNTGGACAAG	499
DB	411	tcacccagctagctcaacaacacatcagatcagagctcggcaaacctctcctgacaag	470
QY	500	TCTTCGGGAACGGAGTCCCTGAATGCCAGCATTTGTGGATGCCATCAACCAAGCTGCTG	559
DB	471	tcttcggggaacgggagtcctgaatgccagctatgtggatgccatcaaccaagctgctg	530
QY	560	ACTGCTGGGTATCCGCTGCCTNCGTTATGAGATCAAGGATATCCATGTGCCACCCCGG	619
DB	531	actgctgggtatccgctgcctcctgttatgagatcaaggatatccatgtgccaccccg	590
QY	620	TGAAGAGTCTATGCAGATCAGGTGAGCGAGCGCGGAAACGGCCACACAGTTCTAG	679
DB	591	tgaagagctctatgcagatgcaggtggaggcagagcgccggaacggccacagctctag	650
QY	680	AGTCTAGGGAGCCCGAGAGTCGCCATCAATGTGGCAGAAGGAGAAACAGAGCCCGAGA	739
DB	651	agtctgaggggaccggagagtcggccatcaatgtggcagaagggaagaaacagggccaga	710
QY	740	TCCTGGCTCCGAAGCAGAAAGGCTGAACAGATAAATCAGGACGAGGAGGAGCCAGTG	799
DB	711	tcttggcctccgaagcagaaagagctgaacagataaaatcaggcagcagagagccagtg	770
QY	800	CAGTTCTGGCGAGGCCAAGCTAAAGCTGAAGCTATTTCGAATCTTGGCTGCAGCTCTGA	859
DB	771	cagttctggcgagggccaaagctaaagctaaagctaatctcgaatcctggctgcaactctga	830
QY	860	CACAACATAATGGAGATGTCAGCAGCTTCACTGACTGTGGCGGAGCAGATATGTTCAGCGCGT	919
DB	831	cacaacataatggagatgagcagcagcttcaactgactgtggccgagcagtatgtcagcgc	890
QY	920	TCTCAAACTGGCCAAAGACTTCCAACTATCTTACTGCTCCCAACCTCCAGCGATGCA	979
DB	891	tctcaaaactggccaaagactcccaactatcctactgcoctccaaacctggcgatgcca	950
QY	980	CCAGCATGGTGGCTCAGGCCATGGGTGTATATGGAGCCCTCACCAGGAGCCAGTGGCAG	1039
DB	951	ccagcatgggtggcctcagggccatgggtgtatatggagccctcaccagagccagtgccgg	1010
QY	1040	GGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGTCCAGGGGTACAGATGCAAGT	1098
DB	1011	ggactccagactcaactctccagtggggagcagcagagatgtccagggtacagatgcaagt	1069
RESULT	8		
AAK51699	ID	AAK51699 standard; cDNA; 1416 BP.	
XX	AC	AAK51699;	
XX	DT	06-NOV-2001 (first entry)	
XX	DE	Human polynucleotide SEQ ID NO 244.	
XX	DE	Human; cytokine; cell proliferation; cell differentiation; gene therapy;	
KW	KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;	
KW	KW	tissue growth factor; immunomodulatory; cancer; leukaemia;	
KW	KW	nervous system disorder; arthritis; inflammation; ss.	
OS	OS	Homo sapiens.	

XX WO200157190-A2.
PN 09-AUG-2001.
XX 05-FEB-2001; 2001WO-US04098.
XX 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0683561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX WPI: 2001-476283/51.
DR P-PSDB; AAM78566.
XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX Claim 1; Page 1135-1136; 6221pp; English.
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX SQ Sequence 1416 BP; 335 A; 364 C; 410 G; 307 T; 0 other;

Query Match 66.8%; Score 793; DB 22; Length 1416;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 893; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 140 GCCGGCTCTCTGGATTGCCCGGAAACACCGTGTACTGTTCGTCGCGAGCAGGAGG 199
Db 150 gccgcctctctgattgcccgaaacacggtgtactgttctgtgcgcagcaggagg 209
Qy 200 CTTGGGTGGTGAGCGAATGGCCGATTCACCCGGATCCTGGAGCCTGTTTGAACATCC 259
Db 210 cctgggtggtgagcgaaatggccgattccaccggtactgagcctggtttgaacatcc 269
Qy 260 TCATCCCTGTGTAGACCGGATCCGATATGTCCAGAGTCTCAGGAAATGTTCATCAACG 319
Db 270 tcatccctgtgtagaccggtatccgatgtgcagagctcgaagaaattgtcatcaacg 329
Qy 320 TGCCTGACAGTCGCTGTGACTCTCGACAATGTAACTCTGCAATCGATGAGTCCCTTT 379
Db 330 tgcctgacagtcggtgtactctcgaacatgttaactctgcaaatcgatggagtccttt 389
Qy 380 ACCTCGCATCATGACCCCTTACAGGCAAGCTACCGGTGTGGAGGACCCCTGAGTATGCCG 439
Db 390 acctgcgcatcatggaaccttacaaggcaagctacggtgtggaggaccctgagtgcg 449
Qy 440 TCACCCAGCTAGCTCAACACCATGATGATCAGACTCGGCNAACCTCTCTNTGGACAAG 499

Db 450 tcaccagctagctcaaacaccatgagatcagctcggaactctctctggcaaaag 509
Qy 500 TCTTCCGGGAACGGAGTCCCTGAATGCCAGCATTTGTGGATGCCATCAACCAAGCTGCTG 559
Db 510 tcttcggggaacgggagtcctgaaatgcagcatgtgtggatgccatcaaccaagctgctg 569
Qy 560 ACTGCTGGGGTATCCGCTGCCTTCGTTATGAGATCAAGGATATCCATGTCACCCCGGG 619
Db 570 actgctggggtatccgctgcctccgttatgagatcaaggatatccatgtgcccacccggg 629
Qy 620 TCAAGAGTCTATGACAGTGCAGGTGGAGGCGAGCGCGGAAACGGGCCACAGTTCCTAG 679
Db 630 tgaagagctctatgagatgcaggtggaggcagagcgcggaacgggcccagttctag 689
Qy 680 AGTCTGAGGGGACCGGAGATGCGGCATCAATGTGGCAGAAGGAAGAAACAGGCCACGA 739
Db 690 agtctgagggggaccgagagtcgcccataatgtggcagaagggaagaaacaggccccaga 749
Qy 740 TCCTGGCTCCGAAGCAGAAAGGCTGAACAGATAAATCAGCAGCAGAGGAGGCCAGTG 799
Db 750 tectggcctccgaagcagaaaagcgtgaacagataaatcagcagcagcagagagccagtg 809
Qy 800 CAGTTCTGCGGAGGCCAAGGCTAAAGCTGAAGCTATTTCGAATCCTGCTGAGCTCTGA 859
Db 810 cagttctggcgaagccaaaggctaaagctgaagctattcgaatcctggtcagctctga 869
Qy 860 CACAACATAATGGAGATGCAGCAGCTTCACCTGCTGTCGCGAGCAGTATGTACGGCGCT 919
Db 870 cacaacataatggagatgcagcagcttcactgacgtgtggccgagcagtatgtcagcgct 929
Qy 920 TCTCCAAACTGSCCAAGGACTCCAACTATCTACTTCCCTCCAACTCCGCGATGTCA 979
Db 930 tctccaaactgsccaaggactccaaactatctactgctcccaacctggcgatgtca 989
Qy 980 CCAGCATGTGCTCAGGCCATGGGTGTATATGAGCCCTCACCAGGCCCCAGT 1034
Db 990 ccagcatggtgctcagccatgggtgtatgtatggagccctcaccacaaagcccaagt 1044

RESULT 9
AAI29477 standard; cDNA; 518 BP.
XX AAI29477;
AC AAI29477;
XX 12-OCT-2001 (first entry)
XX Colon tumour related determined cDNA sequence for clone R0098.F12.
DE Human; immunotherapy; diagnosis; colon cancer; colon tumour;
XX immunogenic; gene therapy; vaccine; colonic cancer; ss.
XX Homo sapiens.
PN WO200149716-A2.
XX 12-JUL-2001.
XX 29-DEC-2000; 2000WO-US35596.
XX 30-DEC-1999; 93US-0476296.
PR 10-JAN-2000; 2000US-0480321.
PR 15-FEB-2000; 2000US-0504629.
PR 06-MAR-2000; 2000US-0519444.
PR 19-MAY-2000; 2000US-0575251.
PR 29-JUN-2000; 2000US-0609448.
PR 28-AUG-2000; 2000US-0649811.
XX (CORI-) CORIXA CORP.
XX Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;
PI King GE, Wang T, Jiang Y;

XX WPI: 2001-441847/47.
XX Colon tumor associated proteins and nucleic acids useful for the
XX prevention, diagnosis and treatment of colonic cancer -
XX
XX Claim 2; Page 415; 472pp; English.
XX
XX The present invention describes colon tumor associated proteins (I) and
XX the polynucleotides (II) that encode them. (I) have cytostatic activity.
XX (I) and (II) can be used in gene therapy and vaccine production. (I) and
XX (II) may be used in the prevention, diagnosis and treatment of diseases
XX associated with inappropriate colon tumor associated protein (TCAP)
XX expression, such as colonic cancer. For example, (I) and (II) may be
XX used to treat disorders associated with decreased expression by
XX rectifying mutations or deletions in a patient's genome that affect the
XX activity of TCAPs by expressing inactive proteins or to supplement the
XX patients own production of them. Additionally, (II) may be used to
XX produce the TCAP proteins, by inserting the nucleic acids into a host
XX cell culturing the cell to express the protein. (II) and its
XX complementary sequences may also be used as DNA probes in diagnostic
XX polymerase chain reaction (PCR) and hybridisation assays to detect and
XX quantitate the presence of similar nucleic acids in samples, and
XX therefore which patients may be in need of restorative therapy. (I) may
XX also be used as antigens in the production of antibodies against TCAPs
XX and in assays to identify modulators of TCAP expression and activity.
XX Anti-(I) antibodies and antagonists may also be used to down regulate
XX TCAP expression and activity. The anti-(I) antibodies may also be used
XX as diagnostic agents for detecting the presence of TCAPs in samples
XX (e.g. by enzyme linked immunosorbant assay (ELISA)). AA128460 to AA129512
XX and AA24494 to AA24523 represent nucleotide and amino acid sequences
XX given in the exemplification of the present invention.
XX
XX Sequence 518 BP: 123 A; 136 C; 149 G; 109 T; 1 other;

Query Match 30.7%; Score 365; DB 22; Length 518;
Best Local Similarity 99.4%; Pred. No. 4.5e-169;
Matches 515; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 200 CCTGGTGGTGGAGCGATGCGGATCCACCGGATCCCTGGAGCCCTGTTGAACATCC 259
DB 1 cctg99t99tggagcgaatg99cgaattccaccggaatcctgagccctg99t99tgaacatcc 60
QY 260 TCATCCCTGTGTTAGCCGGATCGATATGTGCGAGTCTCAAGGAATGTCATCAACG 319
DB 61 tcatccctgtgtagacggaatcgatgtgagagatctcaagaaattgcatcaacg 120
QY 320 TGCTGAGCACTCGGCTGTGACTCTCGACAATGTAACTCTGCAATCGATGGATGCTCTTT 379
DB 121 tgcctgagcagtcg99t99tgaactctgacaaatgtaactctgcaaatgagtgagtccttt 180
QY 380 ACCTGGCGATCATGACCCCTTACAGGCAAGCTACGGTGTGGAGGACCCCTGAGTATGCCG 439
DB 181 acc99cgcatactgagcccttaccaggaacgaagctacg99t99tggaggaacccctgagtcg99 240
QY 440 TCACCCAGCTAGCTCAACAAACCATGATGATCAGAGCTCGGCAAACTCTCTNTGGCAAAAG 499
DB 241 tcacccagctagctcaacaacatgagatgacagctcgcaactctctctgagcaaaag 300
QY 500 TCTTCCGGGAACGGAGTCCCTGAATGCCAGCATGTGAGATGCCATCAACCAAGCTGCTG 559
DB 301 tcttccgggaac999gagtcctggaatg99cagcattgtgagtcgcatcaacccaagctgctg 360
QY 560 ACTGCTGGGTATCCGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 619
DB 361 actgct9999tctccctgctcctcctgctatgagatcaaggaatccatgctg99ccccc999 420
QY 620 TGAAGAGTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 679
DB 421 tgaagagctctatgagatgcatgagtgagagcagcagc999gaaacg99ccacagttctag 480
QY 680 AGTCTGAGGGACCGGAGAGTCGGCCATCAATGTGGCA 717

DB 481 agtctgagggagcccgagagtcggccatcaatgtggca 518
RESULT 10
AAK93925/C
XX ID AAK93925 standard; cDNA; 566 BP.
XX AAK93925;
XX AC AAK93925;
XX DT 06-NOV-2001 (first entry)
XX Human cDNA clone representative sequence, SEQ ID NO: 2385.
XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX Homo sapiens.
XX EP1130094-A2.
XX 05-SEP-2001.
XX 07-JUL-2000; 2000EP-0114089.
XX 08-JUL-1999; 99JP-0194486.
XX 11-JAN-2000; 2000JP-0118774.
XX 02-MAY-2000; 2000JP-0183765.
XX (HELI-) HELIX RES INST.
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI: 2001-524255/58.
XX 830 Primers useful for synthesizing full length cDNA clones and their
XX use in genetic manipulation -
XX Example 11; SEQ ID NO 2385; 1380pp + sequence listing; English.
XX The invention relates to primers for synthesising full length cDNA
XX clones. 830 cDNA molecules encoding a human protein have been
XX isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
XX molecules have been determined. Primers for synthesising the full
XX length cDNA are useful for clarifying the function of the protein encoded by
XX the cDNA. The full length clones were obtained by construction of full
XX length enriched cDNA libraries that were synthesised by the oligo-capping
XX method. The primers enable the production of the full length cDNA easily
XX without any special methods. The present sequence was used as the
XX representative sequence from a human clone
XX homology searches to identify the clone.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in CD-ROM format directly from EPO.
XX Sequence 566 BP: 120 A; 152 C; 139 G; 155 T; 0 other;
Query Match 28.2%; Score 335; DB 22; Length 566;
Best Local Similarity 100.0%; Pred. No. 2.4e-154;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 764 CTGAACAGATAAATCAGCAGCAGGAGGCGGAGTTCGTGGCCGAAGCCCAAGGCTA 823
DB 517 CTGAACAGATAAATCAGCAGCAGGAGGCGGAGTTCGTGGCCGAAGCCCAAGGCTA 458
QY 824 AAGCTGAAGCTATTCGAATCCTGGCTGCAGCTCTGACACACATATGAGATGCAGCAG 883
DB 457 AAGCTGAAGCTATTCGAATCCTGGCTGCAGCTCTGACACACATATGAGATGCAGCAG 398
QY 884 CTTCACTGACTGTGGCGAGCAGTATGTACGCGGTTCTCCAAACTGGCCGAAGGACTCCA 943
DB 397 CTTCACTGACTGTGGCGAGCAGTATGTACGCGGTTCTCCAAACTGGCCGAAGGACTCCA 338

```
QY 944 ACACATATCTACTGCCCTCCAAACCCCTGGCGATGTCACACAGCATGTGGCTCAGGCCATGG 1003
Db 337 ACACATATCTACTGCCCTCCAAACCCCTGGCGATGTCACACAGCATGTGGCTCAGGCCATGG 278
QY 1004 GTGTATATGGAGCCCTCACCRAAGCCCGAGTCCAGGAGCTCCAGACTCCTCCAGTG 1063
Db 277 GTGTATATGGAGCCCTCACCRAAGCCCGAGTCCAGGAGCTCCAGACTCCTCCAGTG 218
QY 1064 GGAGCAGCAGAGATGTCACGAGGTACAGATGCAAGT 1098
Db 217 GGAGCAGCAGAGATGTCACGAGGTACAGATGCAAGT 183

RESULT 11
AAS58752
ID AAS58752 standard; cDNA; 691 BP.
XX
AC AAS58752;
XX
DT 13-FEB-2002 (first entry)
XX
DE cDNA #1428 encoding portion of a human colon tumour protein.
XX
KW Human; colon tumour protein; colon cancer; gene therapy; cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200173027-A2.
XX
PD 04-OCT-2001.
XX
PF 22-MAR-2001; 2001WO-09246.
XX
PR 24-MAR-2000; 2000US-191597P.
PR 04-MAY-2000; 2000US-202024P.
PR 05-MAY-2000; 2000US-202189P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Meagher MJ, Xu J, King GE;
XX
WPI: 2001-611627/70.
XX
New colon tumour proteins and related nucleic acid, useful for
treatment, prevention, diagnosis and monitoring of cancer -
Claim 4; Page 281; 299pp; English.
XX
Th present invention relates to the isolation of novel cDNA sequences
encoding for at least an immunogenic portion of human colon tumour
proteins. The sequences of the invention are useful in pharmaceutical
compositions and vaccines for the prevention and treatment of cancers
such as colon cancer. They are also useful for the diagnosis and
monitoring of such cancers. Antibodies to the colon tumour proteins
and antigen presenting cells that express polynucleotides encoding
colon tumour proteins can be used to inhibit the development of
cancers. T-cells that react specifically with colon tumour proteins
are useful for removing tumour cells from samples (e.g. blood) and
for cancer treatment. The polynucleotides sequences are also useful in
gene therapy. AAS57325-AAS58880 represent the cDNA sequences of the
invention that encode for portions of human colon tumour proteins.
XX
Sequence 691 BP; 183 A; 168 C; 184 G; 140 T; 16 other;
XX
Query Match 27.9%; Score 331; DB 23; Length 691;
Best Local Similarity 99.7%; Pred. No. 2.2e-152;
Matches 381; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 200 CTGGGTGGTGGAGCAATGGCGCATTCACCGGATCCTGGAGCTGCTTTCACATCC 259
Db 1 cctgggtgggagcgaaatgggcccgaattccaccggaatcctggagcctgggttgaacatcc 60
```

```
QY 260 TCATCCCTGTGTAGACCGGATCCGATATGTCAGAGTCTCAAGGAAATGTCATCAACG 319
Db 61 tcatccctgtgtagaccggatccgatattgagagctctcaaggaattgcatcaacg 120
QY 320 TGCCTGAGCAGTCGCTGTGACTCTCGACAATGTAACCTCTGCAAAATCGATGGAGTCTTT 379
Db 121 tgcttgagcagtcgctgtgactctcgacaatgtaactctgcaaatcgatggagtccttt 180
QY 380 ACTCGCGCATCATGACCCCTTACAAGGCAAGCTACCGTGTGGAGGACCTGAGTATGCCG 439
Db 181 acctgcgatcatggaccttacaaggcaagctacggtgtggagacctgagatgcgcg 240
QY 440 TCACCCAGCTAGCTCAACACACATGAGATCAGAGTCGCGCAAACTCTCTNTGGACAAAG 499
Db 241 tcaccagctagctcaaacacacatgagatcagagctcggaactctctctggacaaag 300
QY 500 TCTTCGGGAACGGGAGTCCCTGAATGCCAGCATTTGTGGATGCCATCAACCAAGCTGCTG 559
Db 301 tcttcgggaaaggagtcctcgaaatgccagcattgtggatgccatcaaccaagctgctg 360
QY 560 ACTGCTGGGGTATCCGCTGCCT 581
Db 361 actgctggggtatccgctgcct 382

RESULT 12
AAK94818
ID AAK94818 standard; cDNA; 2064 BP.
XX
AC AAK94818;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human full-length cDNA, SEQ ID NO: 3958.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-0114089.
XX
PR 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
WPI: 2001-524255/58.
DR P-PSDB; AAM93862.
XX
830 Primers useful for synthesizing full length cDNA clones and their
use in genetic manipulation -
PS Claim 8; SEQ ID NO 3958; 1380pp + sequence listing; English.
XX
The invention relates to primers for synthesising full length cDNA
clones. 830 cDNA molecules encoding a human protein have been
isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
molecules have been determined. Primers for synthesising the full length
cDNA are useful for clarifying the function of the protein encoded by
the cDNA. The full length clones were obtained by construction of full
length enriched cDNA libraries that were synthesised by the oligo-capping
method. The primers enable the production of the full length cDNA easily
without any special methods. The present sequence is a full length
human cDNA of the invention.
XX
Note: The sequence data for this patent did not form part of the printed
```

CC specification, but was obtained in CD-ROM format directly from EPO.
 XX
 SQ Sequence 2064 BP: 479 A; 506 C; 581 G; 498 T; 0 other;

Query Match 24.0%; Score 285; DB 22; Length 2064;
 Best Local Similarity 100.0%; Pred. No. 8.4e-130;
 Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 583 CGTTATGAGATCAAGGATATCATGTGCGACCCCGGGTGAAGAGTCTATGCAGATGCAG 642
 |||||
 DB 1169 cgttatgatcaagatatcattgtgccaccgcggtgaaagagtctatgcagatgcag 1228
 |||||
 QY 643 GTGAGGCAGAGCGCGGAACGGCCACAGTCTAGAGTCTGAGGGACCCGAGAGTCG 702
 |||||
 DB 1229 gtggaggcagagcgcggaacggccacagttctagagtctgaggggaccgcgagagtcg 1288
 |||||
 QY 703 GCCATCAATGTGGCAGAAAGGAAGCAAGCCAGATCTCTGGCTCCGAAAGCAGAAAAAG 762
 |||||
 DB 1289 gccatcatgtggcagaagggagaacagagccagatcctggcctccgaagcagaaaaag 1348
 |||||
 QY 763 GCTGAACAGATAAATCAGCGCAGGAGGAGCCAGTGTGAGTCTGTGGCGAAGCCCAAGGCT 822
 |||||
 DB 1349 gctgaacagataaatcaggcagcagagagccagtgagttctgtgcgaagccaaggtc 1408
 |||||
 QY 823 AAAGCTGAAGCTATTCTGAATCTGTGCTCAGCTCTGACACAACT 867
 |||||
 DB 1409 aaagctgaagctattcgaatcctgctgagctctgacacaact 1453
 |||||

RESULT 13
 AAL23495/C
 ID AAL23495 standard; cDNA; 313 BP.
 XX
 AC AAL23495;
 XX
 DT 07-DEC-2001 (first entry)
 XX
 DE Human breast cancer expressed polynucleotide 15952.
 XX
 KW Human; breast cancer; cell marker; cytostatic; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200151628-A2.
 XX
 XX PD 19-JUL-2001.
 XX
 PF 10-JAN-2001; 2001WO-US00798.
 XX
 PR 14-JAN-2000; 2000US-0176077.
 PR 14-MAR-2000; 2000US-0189167.
 PR 24-MAR-2000; 2000US-0192099.
 PR 29-MAR-2000; 2000US-0193480.
 PR 15-MAY-2000; 2000US-0205230.
 PR 09-JUN-2000; 2000US-0211315.
 PR 25-JUL-2000; 2000US-0220534.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Lillie J, Xu Y, Wang Y, Steinmann K;
 XX
 XX WPI; 2001-451856/48.
 XX
 PT New peptide useful as a marker for the diagnosis of breast cancer -
 XX
 PS Claim 1; Page 2911; 3695pp; English.
 XX

The invention relates to human breast cancer expressed polynucleotides (AAL07544-AAL26789) and methods of assessing whether a patient is afflicted with breast cancer by examining the correlation between the expression of certain markers and the cancerous state of breast cells. The polynucleotides and encoded polypeptides are potential markers for

CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity.

XX Sequence 313 BP; 60 A; 86 C; 89 G; 78 T; 0 other;

XX

Query Match	21.3%;	Score 253;	DB 22;	Length 313;
Best Local Similarity	99.7%;	Pred. No. 4.6e-114;		
Matches 303;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy 776	ATCAGGCAGCAGGAGAGGCCAGTCAGTTCCTGCGAAGCCCAAGGCTAAAGCTGAAGCTA	835		
Db 313	ATCAGGCAGCAGGAGAGGCCAGTCAGTTCCTGCGAAGCCCAAGGCTAAAGCTGAAGCTA	254		
Qy 836	TTCAATCTGGCTGCAGCTCTGACACAAACATAATGGAGATGCAGCAGCTTCACGTGACTG	895		
Db 253	TTCAATCTGGCTGCAGCTCTGACACAAACATAATGGAGATGCAGCAGCTTCACGTGACTG	194		
Qy 896	TGGCGGAGCAGTATGTTCAGCGCGTTCCTCAAACTGGCCCAAGGACTCCAACTATCCTTAC	955		
Db 193	TGGCGGAGCAGTATGTTCAGCGCGTTCCTCAAACTGGCCCAAGGACTCCAACTATCCTTAC	134		
Qy 956	TGCCCTCCAAACCCCTGGCGATGTCACACGATGGTGGCTCAGGCCATGGGTGTATATGGAG	1015		
Db 133	TGCCCTCCAAACCCCTGGCGATGTCACACGATGGTGGCTCAGGCCATGGGTGTATATGGAG	74		
Qy 1016	CCCTCACCANAAGCCCCCAGTGCAGGGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGAG	1075		
Db 73	CCCTCACCANAAGCCCCCAGTGCAGGGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGAG	14		
Qy 1076	ATGT 1079			
Db 13	ATGT 10			

RESULT 14
AAS46972/c
ID AAS46972 standard; cDNA; 301 BP.
XX
AC AAS46972;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human breast cancer cDNA clone JBRT21.
XX
KW Human; ss; breast cancer protein; tumour; cancer; cytostatic;
KW gene therapy.
XX
OS Homo sapiens.
XX
PN WO200179286-A2.
XX
PD 25-OCT-2001.
XX
XX 12-APR-2001; 2001WO-US12164.
XX
XX 17-APR-2000; 2000US-0551621.
XX
XX 08-JUN-2000; 2000US-0590751.
XX
XX 22-JUN-2000; 2000US-0604287.
XX
XX 20-JUL-2000; 2000US-0620405.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
XX
XX WPI; 2001-611721/70.
XX
XX Breast Tumour Proteins and nucleic acids useful for the prevention,
PT diagnosis and treatment of breast cancer .
XX
XX Claim 1; Page 155; 297pp; English.
PS

XX The invention relates to isolated breast tumour proteins and
CC nucleic acids that encode them, including immunogenic fragments of the
CC proteins. Also included are expression vectors expressing the
CC proteins, transformed cells and antibodies raised against the proteins or
CC an antigen presenting cell expressing the protein. The proteins and
CC nucleic acids may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate breast tumour protein expression,
CC i.e. breast tumours and breast cancer e.g. by gene therapy. The nucleic
CC acids and their complements may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids in
CC samples, and therefore which patients may be in need of restorative
CC therapy. The proteins, nucleic acids and antibodies may be used in assays
CC to identify modulators (e.g. antagonists) of breast tumour protein
CC expression and activity. The antibodies and antagonists may also be used
CC to down regulate expression and activity. The antibodies may also be used
CC as diagnostic agents for detecting the presence of the proteins in
CC samples (e.g. by enzyme linked immunosorbent assay (ELISA)) and in other
CC immuno-purification diagnostic techniques. The present sequence is
CC a cDNA from a breast tumour cDNA library isolated by subtractive
CC hybridisation against a normal breast cDNA library.
XX
SQ Sequence 301 BP; 48 A; 96 C; 72 G; 85 T; 0 other;

Query Match 21.0%; Score 250; DB 22; Length 301;
Best Local Similarity 99.7%; Pred. No. 1.4e-112;
Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 626 AGTCTATGCAGATCGAGTGGAGCGCGGGAACGGCCACAGTCTTAGAGTCTG 685
DB 301 AGTCTATGCAGATCGAGTGGAGCGCGGGAACGGCCACAGTCTTAGAGTCTG 242
QY 686 AGGGACCCGAGAGTCGCGCCATCAATGTGGCAGAGGGAAGAAACAGGCCCGCAGATCCTGG 745
DB 241 AGGGACCCGAGAGTCGCGCCATCAATGTGGCAGAGGGAAGAAACAGGCCCGCAGATCCTGG 182
QY 746 CCTCGAAGCAGAAAGGCTGAACAGATAAATCAGGCAGCAGGAGAGGCCAGTGCAGTTC 805
DB 181 CCTCGAAGCAGAAAGGCTGCACAGATAAATCAGGCAGCAGGAGAGGCCAGTGCAGTTC 122
QY 806 TGGCGAAGGCCAAGGCTAAAGCTTAAGCTATTTCGAATCCTGGCTGCAGCTCTGCACAAAC 865
DB 121 TGGCGAAGGCCAAGGCTAAAGCTTAAGCTATTTCGAATCCTGGCTGCAGCTCTGCACAAAC 62
QY 866 ATAATGGAGATGCAGAGCTTCACTGACTGTGGCGGAGCAGTATGTGCAGCGGTTCTCCA 925
DB 61 ATAATGGAGATGCAGAGCTTCACTGACTGTGGCGGAGCAGTATGTGCAGCGGTTCTCCA 2
QY 926 A 926
DB 1 A 1

RESULT 15
AAFL17542/c
ID AAF17542 standard; cDNA; 301 BP.
XX
AC AAF17542;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human breast cancer associated JBT721 coding sequence.
XX
KW Human; breast cancer associated gene; vaccine; diagnosis; therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200060076-A2.
XX
PD 12-OCT-2000.
XX
PF 15-FEB-2000; 2000WO-US05308.

XX 02-APR-1999; 99US-0285480.
PR 23-JUN-1999; 99US-0339338.
PR 02-SEP-1999; 99US-0389681.
PR 03-NOV-1999; 99US-0433826.
XX (CORI-) CORIXA CORP.
PA Yuqiu J, Dillon DC, Mitcham JL, Xu J, Harlocker SL;
XX WPI; 2001-122627/13.
DR
XX An isolated polypeptide useful for the treatment and diagnosis of
PT tumors e.g. breast cancer comprises at least an immunogenic portion of
PT a breast tumor protein.
XX
XX Claim 6; Page 95; 238pp; English.
XX The present invention provides the coding sequences and some protein
CC sequences of proteins associated with breast cancer in humans. These
CC sequences can be used in the diagnosis and treatment of cancers,
CC particularly breast tumours.
XX
SQ Sequence 301 BP; 48 A; 96 C; 72 G; 85 T; 0 other;

Query Match 21.0%; Score 250; DB 22; Length 301;
Best Local Similarity 99.7%; Pred. No. 1.4e-112;
Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 626 AGTCTATGCAGATCGAGTGGAGCGCGGGAACGGCCACAGTCTTAGAGTCTG 685
DB 301 AGTCTATGCAGATCGAGTGGAGCGCGGGAACGGCCACAGTCTTAGAGTCTG 242
QY 686 AGGGACCCGAGAGTCGCGCCATCAATGTGGCAGAGGGAAGAAACAGGCCCGCAGATCCTGG 745
DB 241 AGGGACCCGAGAGTCGCGCCATCAATGTGGCAGAGGGAAGAAACAGGCCCGCAGATCCTGG 182
QY 746 CCTCGAAGCAGAAAGGCTGAACAGATAAATCAGGCAGCAGGAGAGGCCAGTGCAGTTC 805
DB 181 CCTCGAAGCAGAAAGGCTGCACAGATAAATCAGGCAGCAGGAGAGGCCAGTGCAGTTC 122
QY 806 TGGCGAAGGCCAAGGCTAAAGCTTAAGCTATTTCGAATCCTGGCTGCAGCTCTGCACAAAC 865
DB 121 TGGCGAAGGCCAAGGCTAAAGCTTAAGCTATTTCGAATCCTGGCTGCAGCTCTGCACAAAC 62
QY 866 ATAATGGAGATGCAGAGCTTCACTGACTGTGGCGGAGCAGTATGTGCAGCGGTTCTCCA 925
DB 61 ATAATGGAGATGCAGAGCTTCACTGACTGTGGCGGAGCAGTATGTGCAGCGGTTCTCCA 2
QY 926 A 926
DB 1 A 1

Search completed: September 22, 2002, 17:25:55
Job time: 4267 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 15:26:33 ; Search time 50.25 Seconds
(without alignments)
5807.216 Million cell updates/sec

Title: US-09-898-216-2
Perfect score: 1188
Sequence: 1 GGCTTCTGGAGCNACCGCT.....GGAAGCAGATTTTCCTGATT 1188

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1188	100.0	1188	1 US-08-781-562-2	Sequence 2, Appli
2	21	1.8	28958	1 US-08-258-261B-6	Sequence 6, Appli
3	21	1.8	28958	1 US-08-456-837-6	Sequence 6, Appli
4	21	1.8	28958	1 US-08-457-342-6	Sequence 6, Appli
5	21	1.8	28958	1 US-08-457-646A-6	Sequence 6, Appli
6	21	1.8	28958	1 US-08-458-076A-6	Sequence 6, Appli
7	21	1.8	28958	1 US-08-764-233A-4	Sequence 4, Appli
8	21	1.8	28958	1 US-08-457-335A-6	Sequence 6, Appli
9	21	1.8	28958	1 US-08-729-214-6	Sequence 6, Appli
10	21	1.8	28958	3 US-09-028-934-6	Sequence 6, Appli
11	21	1.8	49377	1 US-08-764-233A-1	Sequence 6, Appli
12	18	1.5	2371	2 US-08-343-443B-1	Sequence 1, Appli
13	18	1.5	2412	1 US-08-437-027-18	Sequence 18, Appli
14	17	1.4	576	1 US-08-086-428B-34	Sequence 34, Appli
15	17	1.4	576	2 US-08-468-570-34	Sequence 34, Appli
16	17	1.4	576	2 US-08-290-665A-34	Sequence 34, Appli
17	17	1.4	576	5 PCT-US95-10398-34	Sequence 34, Appli
18	17	1.4	603	4 US-09-328-111-244	Sequence 244, App
19	17	1.4	844	4 US-08-998-416-349	Sequence 349, App
20	17	1.4	1025	2 US-08-482-728A-20	Sequence 20, Appli
21	17	1.4	1468	2 US-09-074-512-2	Sequence 2, Appli
22	17	1.4	1721	3 US-09-173-581-16	Sequence 16, Appli
23	17	1.4	1721	4 US-09-420-915-16	Sequence 16, Appli
24	17	1.4	1785	2 US-08-559-505-3	Sequence 3, Appli
25	17	1.4	1785	2 US-08-749-907-3	Sequence 3, Appli
26	17	1.4	1785	4 US-09-241-581B-7	Sequence 7, Appli
27	17	1.4	1785	5 PCT-US95-07721-7	Sequence 7, Appli

28	17	1.4	2060	1 US-08-480-547A-11	Sequence 11, Appl
29	17	1.4	2060	1 US-08-250-847B-11	Sequence 11, Appl
30	17	1.4	2060	2 US-08-463-949A-11	Sequence 11, Appl
31	17	1.4	2060	3 US-08-464-410A-11	Sequence 11, Appl
32	17	1.4	2060	5 PCT-US94-06066-11	Sequence 11, Appl
c 33	17	1.4	2311	4 US-08-976-239-26	Sequence 26, Appli
c 34	17	1.4	35060	3 US-08-814-095-7	Sequence 7, Appli
35	16	1.3	266	1 US-08-594-031-32	Sequence 32, Appli
36	16	1.3	266	1 US-08-594-031-124	Sequence 124, App
37	16	1.3	574	4 US-09-073-297-23	Sequence 23, Appl
38	16	1.3	720	4 US-08-998-416-992	Sequence 992, App
39	16	1.3	876	4 US-08-818-112-19	Sequence 19, Appli
40	16	1.3	876	4 US-08-818-111-19	Sequence 19, Appli
41	16	1.3	876	4 US-09-056-556-19	Sequence 19, Appli
42	16	1.3	977	1 US-08-017-522A-1	Sequence 1, Appli
43	16	1.3	977	6 5215895-2	Patent No. 5215895
44	16	1.3	1018	1 US-08-444-083-7	Sequence 7, Appli
45	16	1.3	1018	1 US-08-286-304-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-08-781-562-2
; Sequence 2, Application US/08781562
; Patent No. 5763589
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN MEMBRANE PROTEIN
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781.562
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/POCKET NUMBER: PF-0181 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1188 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: Consensus
US-08-781-562-2

Query Match 100.0%; Score 1188; DB 1; Length 1188;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	GGCTTCTGGGAGCACCCTCCGCTCGTCTCGTGGTTCGGAGAGTCGCTCGCGGGTGG	60
Db	1	GGCTTCTGGGAGCACCCTCCGCTCGTCTCGTGGTTCGGAGAGTCGCTCGCGGGTGG	60
Qy	61	GAATGCTGGCGGGCGGGCGGGGACACTGGGGCCCTTTTGTGAGGGGCTCTCTACT	120
Db	61	GAATGCTGGCGGGCGGGCGGGGACACTGGGGCCCTTTTGTGAGGGGCTCTCTACT	120
Qy	121	GGCTTCTGGCGGGCTCGGGCGGGCCCTCTCTGATTCGCCGAAACACCGTGTACTG	180
Db	121	GGCTTCTGGCGGGCTCGGGCGGGCCCTCTCTGATTCGCCGAAACACCGTGTACTG	180
Qy	181	TTGCTGCCGACGAGGAGGCTCGGTGTGTGAGGCAATGGGCCGATTCACCGGATCCTG	240
Db	181	TTGCTGCCGACGAGGAGGCTCGGTGTGTGAGGCAATGGGCCGATTCACCGGATCCTG	240
Qy	241	GAGCTGTTTGAACATCCTCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTC	300
Db	241	GAGCTGTTTGAACATCCTCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTC	300
Qy	301	AAGAAATGTTCATCAAGTCGCTGAGCAGTCGGCTGTGACTCTCGACAATGTAAGTCTG	360
Db	301	AAGAAATGTTCATCAAGTCGCTGAGCAGTCGGCTGTGACTCTCGACAATGTAAGTCTG	360
Qy	361	CAATCGATGAGTCTTTACCTCGCATATGAGACCTTACAAGGCAAGCTACGGTGTG	420
Db	361	CAATCGATGAGTCTTTACCTCGCATATGAGACCTTACAAGGCAAGCTACGGTGTG	420
Qy	421	GAGACCTGAGTATGCGCTCACCAGCTAGCTCAACAAACCATGAGATCAGAGTCGGC	480
Db	421	GAGACCTGAGTATGCGCTCACCAGCTAGCTCAACAAACCATGAGATCAGAGTCGGC	480
Qy	481	AACTCTCTNTGGAAAGTCTTCGGGAACGGAGTCCCTGAAATGCCAGCATTTGGAT	540
Db	481	AACTCTCTNTGGAAAGTCTTCGGGAACGGAGTCCCTGAAATGCCAGCATTTGGAT	540
Qy	541	GCCATCAACCAAGTCTGACTGTGGGTATCGCGTCTGCTTCGCTTATGAGATCAAGAT	600
Db	541	GCCATCAACCAAGTCTGACTGTGGGTATCGCGTCTGCTTCGCTTATGAGATCAAGAT	600
Qy	601	ATCATGTGCCACCCGGGTGAAGAGTCTATGAGATGCGAGTGGAGGAGAGCGGGG	660
Db	601	ATCATGTGCCACCCGGGTGAAGAGTCTATGAGATGCGAGTGGAGGAGAGCGGGG	660
Qy	661	AAACGGCCACAGTCTTAGAGTCTGAGGGGACCCGAGAGTCGGCATCAATGTGCGAA	720
Db	661	AAACGGCCACAGTCTTAGAGTCTGAGGGGACCCGAGAGTCGGCATCAATGTGCGAA	720
Qy	721	GGGAAGAACAGGCCCATCTTGGCCCTCCGAAGCAGAAAAAGGCTGAACAGATAATCAG	780
Db	721	GGGAAGAACAGGCCCATCTTGGCCCTCCGAAGCAGAAAAAGGCTGAACAGATAATCAG	780
Qy	781	GCAGCAGAGAGGCGAGTGTGAGTCTGAGAGTCTGAGGGGACCCGAGAGTCGCGCATCAATGTGCGAA	840
Db	781	GCAGCAGAGAGGCGAGTGTGAGTCTGAGAGTCTGAGGGGACCCGAGAGTCGCGCATCAATGTGCGAA	840
Qy	841	ATCTGCTGCTGAGTCTTGACACACAATAATGGAGATGCGAGCTTCACTGACTGTGCC	900
Db	841	ATCTGCTGCTGAGTCTTGACACACAATAATGGAGATGCGAGCTTCACTGACTGTGCC	900
Qy	901	GAGCAGTATGTACGGGCTTCTCCAAACTGGCCAAAGGACTCCAACACTATCTACTGCC	960
Db	901	GAGCAGTATGTACGGGCTTCTCCAAACTGGCCAAAGGACTCCAACACTATCTACTGCC	960
Qy	961	TCCAACTTGGGATGTCACAGCATGGTGTGCTCAGGCCATGGGTGTATATGGAGCCCTC	1020
Db	961	TCCAACTTGGGATGTCACAGCATGGTGTGCTCAGGCCATGGGTGTATATGGAGCCCTC	1020
Qy	1021	ACCAAGCCCCAGTGCACAGGAGTCCAGACTCTCTCTCAGTGGGAGCAGCAGAGATGTC	1080
Db	1021	ACCAAGCCCCAGTGCACAGGAGTCCAGACTCTCTCTCAGTGGGAGCAGCAGAGATGTC	1080

Qy	1081	CAGGTACAGATGCAAGTNTTGTATGAGGAAGTTCGATCGAGTCAAGATGAGTTAGTGGAGC	1140
Db	1081	CAGGTACAGATGCAAGTNTTGTATGAGGAAGTTCGATCGAGTCAAGATGAGTTAGTGGAGC	1140
Qy	1141	TGGGCTTNGCCAGGAGTCTCGGGACAGGAAGAGAGATTTTCTCTGATT	1188
Db	1141	TGGGCTTNGCCAGGAGTCTCGGGACAGGAAGAGAGATTTTCTCTGATT	1188

RESULT 2
US-08-258-261B-6
; Sequence 6, Application US/08258261B
; Patent No. 5639949
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/258,261B
; FILING DATE: 08-JUN-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/457,205
; FILING DATE: 01-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-258-261B-6

Query Match 1.8%; Score 21; DB 1; Length 28958;			
Best Local Similarity 100.0%; Pred. No. 0.46;			
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	129	GCCGCGCTCCGGCGCGCGCTC	149
Db	23859	GCCGCGCTCCGGCGCGCGCTC	23879

```
RESULT 3
US-08-456-837-6
: Sequence 6, Application US/08456837
: Patent No. 5643774
: GENERAL INFORMATION:
: APPLICANT: Schupp, Thomas
: APPLICANT: Ligon, James M.
: APPLICANT: Beck, James Joseph
: APPLICANT: Hill, Dwight Steven
: APPLICANT: Ryals, John Andrew
: APPLICANT: Gaffney, Thomas Deane
: APPLICANT: Lam, Stephen Ting
: APPLICANT: Hammer, Phillip E.
: APPLICANT: Uknes, Scott Joseph
: TITLE OF INVENTION: Genes for the synthesis of
: antipathogenic substances
: NUMBER OF SEQUENCES: 22
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Ciba-Geigy Corporation
: STREET: 7 Skyline Drive
: CITY: Hawthorne
: STATE: NY
: COUNTRY: USA
: ZIP: 10532
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/456.837
: FILING DATE: 01-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/457,205
: FILING DATE: 01-JUN-1995
: APPLICATION NUMBER: 08/258,261
: FILING DATE: 08-Jun-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Elmer, James Scott
: REGISTRATION NUMBER: 36,129
: REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 919-541-8614
: TELEFAX: 919-541-8689
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 28958 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: US-08-456-837-6

Query Match 1.8%; Score 21; DB 1; Length 28958;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 GCCGCGCTCCGCCGCGCTC 149
|||||
Db 23859 GCCGCGCTCCGCCGCGCTC 23879

RESULT 4
US-08-457-342-6
: Sequence 6, Application US/08457342
: Patent No. 5662898
: GENERAL INFORMATION:
: APPLICANT: Schupp, Thomas
```

```
: APPLICANT: Ligon, James M.
: APPLICANT: Beck, James Joseph
: APPLICANT: Hill, Dwight Steven
: APPLICANT: Ryals, John Andrew
: APPLICANT: Gaffney, Thomas Deane
: APPLICANT: Lam, Stephen Ting
: APPLICANT: Hammer, Phillip E.
: APPLICANT: Uknes, Scott Joseph
: TITLE OF INVENTION: Genes for the synthesis of
: antipathogenic substances
: NUMBER OF SEQUENCES: 22
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Ciba-Geigy Corporation
: STREET: 7 Skyline Drive
: CITY: Hawthorne
: STATE: NY
: COUNTRY: USA
: ZIP: 10532
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/457,342
: FILING DATE: 01-JUN-1995
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/457,205
: FILING DATE: 01-JUN-1995
: APPLICATION NUMBER: 08/258,261
: FILING DATE: 08-Jun-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Elmer, James Scott
: REGISTRATION NUMBER: 36,129
: REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 919-541-8614
: TELEFAX: 919-541-8689
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 28958 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: US-08-457-342-6

Query Match 1.8%; Score 21; DB 1; Length 28958;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 GCCGCGCTCCGCCGCGCTC 149
|||||
Db 23859 GCCGCGCTCCGCCGCGCTC 23879

RESULT 5
US-08-457-646A-6
: Sequence 6, Application US/08457646A
: Patent No. 5679560
: GENERAL INFORMATION:
: APPLICANT: Schupp, Thomas
: APPLICANT: Ligon, James M.
: APPLICANT: Beck, James Joseph
: APPLICANT: Hill, Dwight Steven
: APPLICANT: Ryals, John Andrew
: APPLICANT: Gaffney, Thomas Deane
: APPLICANT: Lam, Stephen Ting
: APPLICANT: Hammer, Phillip E.
```

APPLICANT: Uknes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
antipathogenic substances
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,646A
FILING DATE: 01-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-Jun-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-457-646A-6

Query Match 1.8%; Score 21; DB 1; Length 28958;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 129 GCCGGCTCCGGCGCGCCTC 149
Db 23859 GCCGGCTCCGGCGCGCCTC 23879

RESULT 6
US-08-458-076A-6
Sequence 6, Application US/08458076A
Patent No. 5698425
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Gaffney, Thomas Deane
APPLICANT: Lam, Stephen Ting
APPLICANT: Hammer, Phillip E.
APPLICANT: Uknes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
antipathogenic substances
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive

CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,076A
FILING DATE: 01-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-Jun-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-458-076A-6

Query Match 1.8%; Score 21; DB 1; Length 28958;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 129 GCCGGCTCCGGCGCGCCTC 149
Db 23859 GCCGGCTCCGGCGCGCCTC 23879

RESULT 7
US-08-764-233A-4
Sequence 4, Application US/08764233A
Patent No. 5716849
GENERAL INFORMATION:
APPLICANT: Ligon, James M.
APPLICANT: Schupp, Thomas
APPLICANT: Beck, James J.
APPLICANT: Hill, Dwight S.
APPLICANT: Neff, Snezana
APPLICANT: Ryals, John A.
TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/764,233A

;
; FILING DATE: 01-JUN-1995
; APPLICATION NUMBER: 08/258,261
; FILING DATE: 08-Jun-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-457-335A-6

Query Match 1.8%; Score 21; DB 1; Length 28958;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 GCGCGCTCCGCGCGCCTC 149
|||||
DB 23859 GCGCGCTCCGCGCGCCTC 23879

RESULT 8
US-08-457-335A-6
; Sequence 6, Application US/08457335A
; Patent No. 5723759
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammer, Phillip E.
; APPLICANT: Uknes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,335A
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION NUMBER: US 08/457,205

;
; FILING DATE: 01-JUN-1995
; APPLICATION NUMBER: 08/258,261
; FILING DATE: 08-Jun-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-457-335A-6

Query Match 1.8%; Score 21; DB 1; Length 28958;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 GCGCGCTCCGCGCGCCTC 149
|||||
DB 23859 GCGCGCTCCGCGCGCCTC 23879

RESULT 9
US-08-729-214-6
; Sequence 6, Application US/08729214
; Patent No. 5817502
; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Hammer, Phillip E.
; APPLICANT: van Pee, Karl-Heinz
; APPLICANT: Kirner, Sabine
; TITLE OF INVENTION: Genes for the synthesis of
; TITLE OF INVENTION: antipathogenic substances
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 520 White Plains Road
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729,214
; FILING DATE: TBA
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
US-08-729-214-6

Query Match 1.8%; Score 21; DB 1; Length 28958;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 GCCGCGCTCGGCGCGCCTC 149
|||||
DB 23859 GCCGCGCTCGGCGCGCCTC 23879

RESULT 10
US-09-028-934-6
; Sequence 6, Application US/09028934
; Patent No. 6117670
; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Hill, Dwight S.
; APPLICANT: Lam, Steven T.
; APPLICANT: Hammer, Philip E.
; APPLICANT: van Pee, Karl-Heinz
; APPLICANT: Kirner, Sabine
; APPLICANT: Young, Thomas R.
; TITLE OF INVENTION: Pyrrolnitrin Biosynthesis Genes and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6117670artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/028.934
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/729,214
; FILING DATE: 09-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/258,261
; FILING DATE: 08-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 36,241
; REFERENCE/DOCKET NUMBER: CGC1506/CIP7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-09-028-934-6

Query Match 1.8%; Score 21; DB 3; Length 28958;

Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 GCCGCGCTCGGCGCGCCTC 149
|||||
DB 23859 GCCGCGCTCGGCGCGCCTC 23879

RESULT 11
US-08-764-233A-1
; Sequence 1, Application US/08764233A
; Patent No. 5716849
; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Schupp, Thomas
; APPLICANT: Beck, James J.
; APPLICANT: Hill, Dwight S.
; APPLICANT: Neff, Snezana
; APPLICANT: Ryals, John A.
; TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 520 White Plains Road, P.O. Box 2005
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/764.233A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/729,214
; FILING DATE: 09-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/258,261
; FILING DATE: 08-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 36,241
; REFERENCE/DOCKET NUMBER: 1506/CIP6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49377 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Sorangium cellulosum
; IMMEDIATE SOURCE:
; CLONE: p98/1, pJL3, and pVRM15
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 383..760
; OTHER INFORMATION: /product= "SorR"
; OTHER INFORMATION: /note= "This gene encodes a protein that is highly homolog
; OTHER INFORMATION: the reductase domains of type I PKSs such as eryA from
; OTHER INFORMATION: Saccharopolyspora erythraea."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 927..19874
; OTHER INFORMATION: /product= "SorA"
; OTHER INFORMATION: /note= "Gene product is highly homologous to type I PKSs t

OTHER INFORMATION: are known to be involved in the synthesis of polyketide
OTHER INFORMATION: compounds."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 942..7115
OTHER INFORMATION: /product= "Module 1 of SorA"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7203..12884
OTHER INFORMATION: /product= "Module 2 of SorA"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 13455..19616
OTHER INFORMATION: /product= "Module 3 of SorA"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19871..46318
OTHER INFORMATION: /product= "SorB"
OTHER INFORMATION: /note= "Gene product is highly homologous to type I PKS genes"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19870..24556
OTHER INFORMATION: /product= "Module 1 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 24638..30820
OTHER INFORMATION: /product= "Module 2 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 30881..35446
OTHER INFORMATION: /product= "Module 3 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 35528..40114
OTHER INFORMATION: /product= "Module 4 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 40190..46318
OTHER INFORMATION: /product= "Module 5 of SorB"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 46851..47891
OTHER INFORMATION: /product= "Sorm"
OTHER INFORMATION: /note= "The protein encoded by the sorm gene is highly homologous to the methyltransferase from Streptomyces hygroscopicus that is involved in the synthesis of the polyketide rapamycin."
US-08-764-233A-1

Query Match 1.8%; Score 21; DB 1; Length 49377;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 GCCGGCTCCGGCGGCCTC 149
|||||
Db 40162 GCCGGCTCCGGCGGCCTC 40182

RESULT 12
US-08-343-443B-1/c
Sequence 1, Application US/08343443B
Patent No. 5968734
GENERAL INFORMATION:
APPLICANT: Aurias, Alain
APPLICANT: Delattre, Olivier
APPLICANT: Desmaze, Chantal
APPLICANT: Melot, Thomas
APPLICANT: Peter, Martine
APPLICANT: Plooungastel, Beatrice
APPLICANT: Thomas, Gilles
APPLICANT: Zucman, Jessica
TITLE OF INVENTION: NUCLEIC ACID CORRESPONDING TO A GENE OF

TITLE OF INVENTION: CHROMOSOME 22 INVOLVED IN RECURRENT CHROMOSOMAL
TITLE OF INVENTION: TRANSLATIONS ASSOCIATED WITH THE DEVELOPMENT OF CANCEROUS
TITLE OF INVENTION: TUMORS, AND NUCLEIC ACIDS OF FUSION RESULTING FROM SAID
TITLE OF INVENTION: TRANSLOCATIONS
NUMBER OF SEQUENCES: 129
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weiser & Associates
STREET: 230 South Fifteenth Street
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: AEDIT 1.0 DOS text editor
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/343,443B
FILING DATE: 18-NOV-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/00494
FILING DATE: 19-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/06123
FILING DATE: 20-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Weiser, Gerard J.
REGISTRATION NUMBER: 19,763
REFERENCE/DOCKET NUMBER: 989.6121P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-875-8383
TELEFAX: 215-875-8394
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2371 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 25..1992
US-08-343-443B-1

Query Match 1.5%; Score 18; DB 2; Length 2371;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 AGGAGCCCTGGGTGGTG 211
|||||
Db 358 AGGAGCCCTGGGTGGTG 341

RESULT 13
US-08-437-027-18/c
Sequence 18, Application US/08437027
Patent No. 5670317
GENERAL INFORMATION:
APPLICANT: Landanyi, Marc
APPLICANT: Gerald, William
TITLE OF INVENTION: A DIAGNOSTIC TEST FOR TEST FOR THE DESMOPLASTIC
TITLE OF INVENTION: SMALL ROUND CELL TUMOR
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: US/08/437,027
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 46416/JPM/CCA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 2412 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-437-027-18

Query Match 1.5%; Score 18; DB 1; Length 2412;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 AGGAGCGCTGGTGGTGG 211

DB 358 AGGAGCGCTGGTGGTGG 341

RESULT 14

US-08-086-428B-34
Sequence 34, Application US/08086428B
Patent No. 5514539
GENERAL INFORMATION:

APPLICANT: BUKH, J., MILLER, R.H. AND
PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
OF 51 ISOLATES OF HEPATITIS C AND THE USE
OF REAGENTS DERIVED FROM THESE SEQUENCES IN
DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 159
CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/086,428B
FILING DATE: 29-JUN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4070
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792

INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:

LENGTH: 576 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
ORIGINAL SOURCE:
ORGANISM: homosapiens
INDIVIDUAL ISOLATE: S83
US-08-086-428B-34

Query Match 1.4%; Score 17; DB 1; Length 576;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 554 CTGCTGACTGCTGGGGT 570

DB 553 CTGCTGACTGCTGGGGT 569

RESULT 15

US-08-468-570-34
Sequence 34, Application US/08468570
Patent No. 5871962
GENERAL INFORMATION:

APPLICANT: BUKH, J., MILLER, R.H. AND
PURCELL, R.H.
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
OF 51 ISOLATES OF HEPATITIS C AND THE USE
OF REAGENTS DERIVED FROM THESE SEQUENCES IN
DIAGNOSTIC METHODS AND VACCINES
NUMBER OF SEQUENCES: 159
CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/468,570
FILING DATE: 6-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/086,428
FILING DATE: 29-JUN-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4070US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792

INFORMATION FOR SEQ ID NO: 34:

SEQUENCE CHARACTERISTICS:
LENGTH: 576 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
ORIGINAL SOURCE:
ORGANISM: homosapiens
INDIVIDUAL ISOLATE: S83
US-08-468-570-34

Query Match

1.4%; Score 17; DB 2; Length 576;

Matches 639; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 264 CCCTGTGTAGACCGGATCCGATATGTGACAGAGTCTCAAGAAATGTGCATCAACGTGCC 323
Db 2 CCCTGTGTAGACCGGATCCGATATGTGACAGAGTCTCAAGAAATGTGCATCAACGTGCC 61

QY 324 TGACAGTGGCTGTGACTCTGACATGTACTCTGCAATTCGATGGAGTCTTTACCT 383
Db 62 TGACAGTGGCTGTGACTCTGACATGTACTCTGCAATTCGATGGAGTCTTTACCT 121

QY 384 GCGCATCATGAGCCCTTCAAGGCAAGCTAGCGTGTGAGGACCTGAGTATGCCGTAC 443
Db 122 GCGCATCATGAGCCCTTCAAGGCAAGCTAGCGTGTGAGGACCTGAGTATGCCGTAC 181

QY 444 CCAGCTAGCTCAAAACCAACCATGAGATCAGAGCTGGCAAACTCTGCTGCAAAAGTCTT 503
Db 182 CCAGCTAGCTCAAAACCAACCATGAGATCAGAGCTGGCAAACTCTGCTGCAAAAGTCTT 241

QY 504 CCGGGAACGGAGTCCCTGGAATGCCAGCTTGTGATGCCATCAACAGAGTCTGACTG 563
Db 242 CCGGGAACGGAGTCCCTGGAATGCCAGCTTGTGATGCCATCAACAGAGTCTGACTG 301

QY 564 CTGGGATATCCGCTGCGCTTCGCTTATGAGATCAAGGATATCATGTCGCCACCCGGGTGA 623
Db 302 CTGGGATATCCGCTGCGCTTCGCTTATGAGATCAAGGATATCATGTCGCCACCCGGGTGA 361

QY 624 AGACTATATGAGATGAGTGAAGGAGAGAGCGGCAAGGAGTCTTGAAGTCT 683
Db 362 AGACTATATGAGATGAGTGAAGGAGAGAGCGGCAAGGAGTCTTGAAGTCT 421

QY 684 TGAGGGAGCCGAGAGTGGCCATCATGTGCGAGAGAGAGAAACAGGCCAGATCTT 743
Db 422 TGAGGGAGCCGAGAGTGGCCATCATGTGCGAGAGAGAGAAACAGGCCAGATCTT 481

QY 744 GGCCTCCGAAGCAGAAAAAGGCTGAACAGATCAAGCAGCAGAGAGAGGCGCAGTGCAGT 803
Db 482 GGCCTCCGAAGCAGAAAAAGGCTGAACAGATCAAGCAGCAGAGAGAGGCGCAGTGCAGT 541

QY 804 TCTGGCAGAGCCCAAGGCTTAAGCTGAAGCTATTGCAATCTGCGCTGAGCTCTACACA 863
Db 542 TCTGGCAGAGCCCAAGGCTTAAGCTGAAGCTATTGCAATCTGCGCTGAGCTCTGAGACA 601

QY 864 ACATATATGAGATGAGCAGAGCTTCACTGACTGTGGCGAGC 904
Db 602 ACATATATGAGATGAGCAGAGCTTCACTGACTGTGGCGAGC 642

RESULT 15
BM471641 1037 bp mRNA Linear EST 05-FEB-2002
LOCUS AGENCOURT_6465267 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5539088
DEFINITION 5', mRNA sequence.
ACCESSION BM471641
VERSION BM471641.1 GI:18520683
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1037)
NIH-MGC <http://mgc.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph. D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC/DCID/DBP
CDNA Library Preparation: Life Technologies Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LHAM12233 row: a column: 09

High quality sequence stop: 589.
FEATURES
source
Location/Qualifiers
1..1037
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5539088"
/clone_id="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

BASE COUNT 236 a 321 c 251 g 229 t
ORIGIN

Query Match 45.1%; Score 536; DB 10; Length 1037;
Best Local Similarity 99.7%; Pred. No. 3.7e-259;
Matches 636; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 175 GTACGTTCGTCGCGCAGAGAGGCGTGGTGTGAGAGCCGAATGGCCGATTCACCGG 234
Db 1 GTACGTTCGTCGCGCAGAGAGGCGTGGTGTGAGAGCCGAATGGCCGATTCACCGG 60

QY 235 ATCTGAGACCTGCTTTGACATCCTCATCCCTGTTAGACCGGATCCGATATGTGCAG 294
Db 61 ATCTGAGACCTGCTTTGACATCCTCATCCCTGTTAGACCGGATCCGATATGTGCAG 120

QY 295 AGTCTCAAGAAATTTGTCATCAAGCTGCTGAGCAGAGTGGCTGTGACTCTGACAAATGA 354
Db 121 AGTCTCAAGAAATTTGTCATCAAGCTGCTGAGCAGAGTGGCTGTGACTCTGACAAATGA 180

QY 355 ACTCTGCAAAATCGATGAGTGCCTTTACTGGGCAATCAAGACCCCTTAAAGGCAAGCTAC 414
Db 181 ACTCTGCAAAATCGATGAGTGCCTTTACTGGGCAATCAAGACCCCTTAAAGGCAAGCTAC 240

QY 415 GGTGTGAGAGACCTGAGATATGCGTACCCAGCAGTAGTCAAAACCAACATGAGATCAGAG 474
Db 241 GGTGTGAGAGACCTGAGATATGCGTACCCAGCAGTAGTCAAAACCAACATGAGATCAGAG 300

QY 475 CTGGGCAAACTCTCTNTGCAAAAGTCTTCGGGAACGGAGTCCCTGAATGCCAGATT 534
Db 301 CTGGGCAAACTCTCTNTGCAAAAGTCTTCGGGAACGGAGTCCCTGAATGCCAGATT 360

QY 535 GTGATGCCATCAACCAAGCTGCTGACTGTGGGATATCCCTGCTTCGTTATGAGATC 594
Db 361 GTGATGCCATCAACCAAGCTGCTGACTGTGGGATATCCCTGCTTCGTTATGAGATC 420

QY 595 AAGATATCCATGTGCGCACCCCGGCTGAAGAAGTCTATGACAGATGAGGTGAGAGAGAG 654
Db 421 AAGATATCCATGTGCGCACCCCGGCTGAAGAAGTCTATGACAGATGAGGTGAGAGAGAG 480

QY 655 CGCGGAAACGGGCGACAGTTCTAGAGTCTGAGGGAGACCCGAGAGTGGCCATCAATGTG 714
Db 481 CGCGGAAACGGGCGACAGTTCTAGAGTCTGAGGGAGACCCGAGAGTGGCCATCAATGTG 540

QY 715 GCAAGAAGGAGAAACAGCGCCCAATCTGCGCTCCCAAGAGAAAAAGGCTGAACAGATA 774
Db 541 GCAAGAAGGAGAAACAGCGCCCAATCTGCGCTCCCAAGAGAAAAAGGCTGAACAGATA 600

QY 775 AATCAGCAGCAGAGAGAGGCGCAGTGTGCGGAA 812
Db 601 AATCAGCAGCAGAGAGAGGCGCAGTGTGCGGAA 638

Search completed: September 22, 2002, 16:48:50
Job time: 5157 sec

```
Db 601 GCAGTATGTACCGCGCTTCCTCAAACTGGCCAGAGCTCCAACTCACTACTGCTCC 657
|||||
RESULT 13
LOCUS BG036463 749 bp mRNA linear EST 24-JAN-2001
DEFINITION 60232409F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4427787 5',
mRNA sequence.
ACCESSION BG036463
VERSION BG036463.1 GI:12431652
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 749)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM10177 row: a column: 04
High quality sequence stop: 708.
Location/Qualifiers
1. 749
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4427787"
/clone_lib="NIH_MGC_91"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: prostate; Vector: pCMV-SPORT6; site_1: NotI;
site_2: SalI; Cloned unidirectionally: oligo-dT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH-MGC Library."
BASE COUNT 190 a 183 c 222 g 154 t
ORIGIN
Query Match 46.5%; Score 552; DB 10; Length 749;
Best Local Similarity 99.8%; Pred. No. 2.9e-267;
Matches 602; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

496 AAGGTTCCGGGAGCGGAGTCCGTAATGCGAGATTGTGATCCATCAACGAAGCT 555
|||||
Db 1 AAGGTTCCGGGAGCGGAGTCCGTAATGCGAGATTGTGATCCATCAACGAAGCT 60
|||||
556 GCTGACTGCTGGGGTATCCGCTGCTTCGTTATGAGATCAAGATATCCATGTCACCC 615
|||||
Db 61 GCTGACTGCTGGGGTATCCGCTGCTTCGTTATGAGATCAAGATATCCATGTCACCC 120
|||||
616 CGGGTAAAGAGCTATGATGATGATGAGGAGGAGGCGGGAAGCGGCACAGATT 675
|||||
Db 121 CGGGTAAAGAGCTATGATGATGATGAGGAGGAGGCGGGAAGCGGCACAGATT 180
|||||
676 CTGAGCTCTGAGGAGCCGAGAGTGGCCATCAATGTGGCAGAGGAGGAAGAGGCC 735
|||||
Db 181 CTGAGCTCTGAGGAGCCGAGAGTGGCCATCAATGTGGCAGAGGAGGAAGAGGCC 240
|||||
736 CAGATCCTGGCTCTCCAGAGCAAAAAGCTGAACAGATAATCAGGACGAGAGAGGCC 795
|||||
Db 241 CAGATCCTGGCTCTCCAGAGCAAAAAGCTGAACAGATAATCAGGACGAGAGAGGCC 300
|||||
796 AGTGCAGTTCTGGCGAAGGCCAAGGCTAAGAGCTATGGAATCCTGGCTGACGCT 855
|||||
```

```
Db 301 AGTGAGTTCTGGCGAAGGCCAAGGCTAAGCTAATTCATCTGGCTGACACT 360
|||||
856 CTGACACAACATTAATGAGATGACAGCTTCACATGATGAGCGAGACAGATGTACAGC 915
|||||
Db 361 CTGACACAACATTAATGAGATGACAGCTTCACATGATGAGCGAGACAGATGTACAGC 420
|||||
916 GCGTTCTCCAAACTGGCCAGAGACTCCAACTATCTCACTGCTCCCAACCTGGCGAT 975
|||||
Db 421 GCGTTCTCCAAACTGGCCAGAGACTCCAACTATCTCACTGCTCCCAACCTGGCGAT 480
|||||
976 GTACACAGATGATGGCTTCAGGCCATGGGTGTATATGAGAGCCCTCAACAGCCGAGTG 1035
|||||
Db 481 GTACACAGATGATGGCTTCAGGCCATGGGTGTATATGAGAGCCCTCAACAGCCGAGTG 540
|||||
1036 CCAGGACTCCAGACTCTCTCCAGTGGAGACAGACAGATGTCCAGAGGTACAGATGCA 1095
|||||
Db 541 CCAGGACTCCAGACTCTCTCCAGTGGAGACAGACAGATGTCCAGAGGTACAGATGCA 600
|||||
1096 AGT 1098
|||
Db 601 AGT 603
|||||
RESULT 14
LOCUS BG828070 770 bp mRNA linear EST 22-MAY-2001
DEFINITION 60275343F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4905945 5',
mRNA sequence.
ACCESSION BG828070
VERSION BG828070.1 GI:14175657
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 770)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM1808 row: d column: 10
High quality sequence stop: 767.
Location/Qualifiers
1. 770
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4905945"
/clone_lib="NIH_MGC_17"
/tissue_type="thabdomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: muscle; Vector: pOTB7; site_1: EcoRI;
site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(C). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the Laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

```
BASE COUNT 200 a 206 c 216 g 148 t
ORIGIN
Query Match 45.4%; Score 539; DB 10; Length 770;
Best Local Similarity 99.7%; Pred. No. 1.1e-260;
```



```

OY 800 CAGTTGCGGCAAGGCCAAGCTAAAGCTGATTCGATCCTGGCTCAGCTCTGA 859
      |||
Db 791 CAGTTGCGGCAAGGCCAAGCTAAAGCTGATTCGATCCTGGCTCAGCTCTGA 850
OY 860 CACA 863
      |||
Db 851 CACA 854

RESULT 10
LOCUS B1092184 796 bp mRNA linear EST 20-JUN-2001
DEFINITION 602856421F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:4997837 5',
            mRNA sequence.
ACCESSION B1092184
VERSION B1092184.1 GI:14510514
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-remail.nih.gov
          Tissue Procurement: ATCC
          CDNA Library Preparation: Life Technologies, Inc.
          CDNA Library Arrayed by: Incyte Genomics, Inc.
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LHAM1025 row: i column: 06
          High quality sequence stop: 793.
          Location/Qualifiers
            1. 796
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="IMAGE:4997837"
               /clone_lib="NIH_MGC_10"
               /cell_line="MGC36"
               /lab_host="DH10B"
               /note="Organ: cervix; Vector: PCMV-SPORT6; Site:1: NotI;
               Site:2: SalI; Cloned unidirectionally. Primer: oligo dt.
               Average insert size 1.5 kb. Library prepared by Life
               Technologies."
BASE COUNT 206 a 214 c 230 g 146 t
ORIGIN
Query Match 47.88; Score 568; DB 10; Length 796;
Best Local Similarity 99.78; Pred. No. 2.5e-275;
Matches 668; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 352 GTAACCTGCAAAATGATGAGATCCTTACCTGCGCATCATGAGCCCTTACAAGCAAGC 411
      |||
Db 42 GTAACCTGCAAAATGATGAGATCCTTACCTGCGCATCATGAGCCCTTACAAGCAAGC 101
OY 412 TAGCGGTGAGAGACCCGTAGTATCCCTACCCAGCACTAGCTCAAAACAACATGAGATCA 471
      |||
Db 102 TAGCGGTGAGAGACCCGTAGTATCCCTACCCAGCACTAGCTCAAAACAACATGAGATCA 161
OY 472 GAGCTCGGCAAACTCTCTNTGGACAAAGCTTCGGGAAAGGAGATCCCTGAATGCCAGC 531
      |||
Db 162 GAGCTCGGCAAACTCTCTNTGGACAAAGCTTCGGGAAAGGAGATCCCTGAATGCCAGC 221
OY 532 ATTGTGATGCCATCAACCAAGCTGCTGACTGCTGGGTATCCGCTGCTTATGAG 591
      |||
Db 222 ATTGTGATGCCATCAACCAAGCTGCTGACTGCTGGGTATCCGCTGCTTATGAG 281
OY 592 ATCAAGGATATCATGTGCGCACCCCGGGTGAAGAAGTCTATGACATGACAGGTGAGGCA 651

```

```

      |||
Db 282 ATCAAGGATATCATGTGCGCACCCCGGGTGAAGAAGTCTATGACATGACAGTGGAGGCA 341
OY 652 GAGCGGGGAAAGCGGCGACAGTTCTAGAGTCTGAGGGGACCCGAGATCCGGCATCAAT 711
      |||
Db 342 GAGCGGGGAAAGCGGCGACAGTTCTAGAGTCTGAGGGGACCCGAGATCCGGCATCAAT 401
OY 712 GTGCGAGAAAGGAAAGACAGAGCCAGATCCTGGCTCCGAAGCAGAAAGAGGCTGACAG 771
      |||
Db 402 GTGCGAGAAAGGAAAGACAGAGCCAGATCCTGGCTCCGAAGCAGAAAGAGGCTGACAG 461
OY 462 ATAAATCAGCAGCAGAGAGAGGCCAGTGTGCGGAAAGGCCAAGCTAAAGCTGAA 521
OY 832 GCTATTCGAATCTCTGCGTCTGACAGCTCTGACACAACATATGAGAGATGACAGCTTCACTG 891
      |||
Db 522 GCTATTCGAATCTCTGCGTCTGACAGCTCTGACACAACATATGAGAGATGACAGCTTCACTG 581
OY 892 ACTGTGCGCGAGCAGATGATGACGCGCTTCACAAACTGGCCAAAGACTCCAACTATC 951
      |||
Db 582 ACTGTGCGCGAGCAGATGATGACGCGCTTCACAAACTGGCCAAAGACTCCAACTATC 641
OY 952 CTACTGCCCTCCAAACCCGTGGGATGTCACACAGATGTGCTCAGGCCATGGGTATAT 1011
      |||
Db 642 CTACTGCCCTCCAAACCCGTGGGATGTCACACAGATGTGCTCAGGCCATGGGTATAT 701
OY 1012 GGAGCCCTCA 1021
      |||
Db 702 GGAGCCCTCA 711

```

```

RESULT 11
LOCUS BE740453 660 bp mRNA linear EST 15-SEP-2000
DEFINITION 601595494F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3949582 5',
            mRNA sequence.
ACCESSION BE740453
VERSION BE740453.1 GI:10154445
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-remail.nih.gov
          Tissue Procurement: DCTD/DHP
          CDNA Library Preparation: Ling Hong/Rubin Laboratory
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
          Plate: LUCM814 row: k column: 23
          High quality sequence stop: 660.
          Location/Qualifiers
            1. 660
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="IMAGE:3949582"
               /clone_lib="NIH_MGC_9"
               /tissue_type="adenocarcinoma cell line"
               /lab_host="DH10B (phage-resistant)"
               /note="Organ: ovary; Vector: pOTB7; Site:1: XhoI; Site:2:
               EcoRI; cDNA made by oligo-dt priming. Directionally
               cloned into EcoRI/XhoI sites using the following 5'
               adaptor: GGCACGAG(G). Size-selected >500bp for average
               insert size 1.8kb. Library constructed by Ling Hong in
               the laboratory of Gerald M. Rubin (University of
               California, Berkeley) using ZAP-cDNA synthesis kit

```

Query Match 48.1% Score 572; DB 10; Length 887;
 Best Local Similarity 99.7% Pred. No. 2,5e-277;
 Matches 672; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

252 GAACATCCCATCCCTGTTAGACCGGATCGGATGTGTGACAGTCTCAGGAATTTGT 311
 |||||||
 2 GAACATCCCATCCCTGTTAGACCGGATCGGATGTGTGACAGTCTCAGGAATTTGT 61

312 CATCAACGTCCTGAGACAGTCCGCTGTGATCTCGACAAATGTAATCTCAATCATG 371
 |||||||
 62 CATCAACGTCCTGAGACAGTCCGCTGTGATCTCGACAAATGTAATCTCAATCATG 121

372 AGTCCTTTACCTGCGCATATGAGACCTTTACAAGCAAGCTACGCTGTGAGAGACCTGA 431
 |||||||
 122 AGTCCTTTACCTGCGCATATGAGACCTTTACAAGCAAGCTACGCTGTGAGAGACCTGA 181

432 GTATGCGCGTACCGACATGCTCAACCAACATGATGAGAGCTCGCAACTCTCTMT 491
 |||||||
 182 GTATGCGCGTACCGACATGCTCAACCAACATGATGAGAGCTCGCAACTCTCTMT 241

492 GGACAAAGTCTTCCGGGAGAGGGAGTCCCTGATGCGCAGCATTTGTGATGCCATCAACCA 551
 |||||||
 242 GGACAAAGTCTTCCGGGAGAGGGAGTCCCTGATGCGCAGCATTTGTGATGCCATCAACCA 301

552 AGCTCTGACTCTGCGGATATCCCTGCTTNCCTTATGATCAAGATATCCATGCTCC 611
 |||||||
 302 AGCTCTGACTCTGCGGATATCCCTGCTTNCCTTATGATCAAGATATCCATGCTCC 361

612 ACCCGGGTGAAGAGTCTATGAGATGAGGTGGAGGAGGCGGAGAAAGGCGCCAC 671
 |||||||
 362 ACCCGGGTGAAGAGTCTATGAGATGAGGTGGAGGAGGCGGAGAAAGGCGCCAC 421

672 AGTTTGAAGTCTGAGGGAGCCGAGAGTGGCCATCAATGTGCGAGAAAGGAGAAACA 731
 |||||||
 422 AGTTTGAAGTCTGAGGGAGCCGAGAGTGGCCATCAATGTGCGAGAAAGGAGAAACA 481

732 GGGCCAGATCTTGGGCTCCGAGAGCAAAAAGCTGAACAGATTAATCAGGCGCAGAGAA 791
 |||||||
 482 GGGCCAGATCTTGGGCTCCGAGAGCAAAAAGCTGAACAGATTAATCAGGCGCAGAGAA 541

792 GGGCCAGTCAAGTCTGCGGAGAGGCAAGGCTAAAGCTGAAGTATTGCAATCTGGCTGC 851
 |||||||
 542 GGGCCAGTCAAGTCTGCGGAGAGGCAAGGCTAAAGCTGAAGTATTGCAATCTGGCTGC 601

852 AGCTTGACACAAATATGAGATGACAGACAGCTTCACTGACTGTGGCCGAGCAGATGT 911
 |||||||
 602 AGCTTGACACAAATATGAGATGACAGACAGCTTCACTGACTGTGGCCGAGCAGATGT 661

912 CAGCGGTTTCCCA 925
 |||||||
 662 CAGCGGTTTCCCA 675

RESULT 9
 BM476304 1110 bp mRNA linear EST 05-FEB-2002
 LOCUS BM476304
 DEFINITION AGNCOURT_6479118 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5558923
 5', mRNA sequence.
 ACCESSION BM476304
 VERSION BM476304.1 GI:18525346
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1110)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.

cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LHAM1282 row: k column: 20
 High quality sequence start: 8
 High quality sequence stop: 678.
 Location/Qualifiers
 1. 1110
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5558923"
 /clone_11b="NIH_MGC_88"
 /tissue_type="quodonal adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: small intestine; Vector: pcMV-SpO6;
 Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
 oligo-dr primed. Average insert size 1.767 kb. Library
 enriched for full-length clones and constructed by Life
 Technologies. Note: this is a NIH_MGC Library."
 BASE COUNT 249 a 322 c 326 g 213 t

Query Match 48.1% Score 571; DB 10; Length 1110;
 Best Local Similarity 99.6% Pred. No. 8.4e-277;
 Matches 721; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

140 GCCGCGCTCTCTGATTTGCCCGCAACACCGTGTACTGTTGCGCGCAGCAGAGAG 199
 |||||||
 131 GCCGCGCTCTCTGATTTGCCCGCAACACCGTGTACTGTTGCGCGCAGCAGAGAG 190

200 CCTGGTGTGAGGCAATGAGGCGCATTCACCGGATCTTGAGGCTGGTTGAACATCC 259
 |||||||
 191 CCTGGTGTGAGGCAATGAGGCGCATTCACCGGATCTTGAGGCTGGTTGAACATCC 250

260 TCATCCCTGTGTTAGACCGGATCCGATATGTGCAGAGTCTCAAGAAATTTGCATCAACG 319
 |||||||
 251 TCATCCCTGTGTTAGACCGGATCCGATATGTGCAGAGTCTCAAGAAATTTGCATCAACG 310

320 TGCCTGAGAGTCTGCTGTGATCTCGAATATGTAATCTGCAATTCATGAGTCTTT 379
 |||||||
 311 TGCCTGAGAGTCTGCTGTGATCTCGAATATGTAATCTGCAATTCATGAGTCTTT 370

380 ACCTGGCATCATGAGACCTTACAAGGCAAGCTAGCGGTGAGAGACCTGATGATGCCG 439
 |||||||
 371 ACCTGGCATCATGAGACCTTACAAGGCAAGCTAGCGGTGAGAGACCTGATGATGCCG 430

440 TCACCCAGCTAGCTCAACACCATGAGATGATGAGCTCGGCAAACTCTGTTGACCAAG 499
 |||||||
 431 TCACCCAGCTAGCTCAACACCATGAGATGATGAGCTCGGCAAACTCTGTTGACCAAG 490

500 TCTTCCGGGAAGGAGAGTCCCTGATGATGATGATGATGATGATGATGATGATGATG 559
 |||||||
 491 TCTTCCGGGAAGGAGAGTCCCTGATGATGATGATGATGATGATGATGATGATGATG 550

560 ACTGCTGGGATGTCGCGTCCCTGATGATGATGATGATGATGATGATGATGATGATG 619
 |||||||
 551 ACTGCTGGGATGTCGCGTCCCTGATGATGATGATGATGATGATGATGATGATGATG 610

620 TGAAGAGTCTATGAGATGAGTGGAGAGCAGAGCGGCGGAAAGCGGCGACAGTTTGA 679
 |||||||
 611 TGAAGAGTCTATGAGATGAGTGGAGAGCAGAGCGGCGGAAAGCGGCGACAGTTTGA 670

680 AGTCTGAGGAGACCCGAGAGTGGCCATCATATGTGCGAGAGGAGAAAGAGGCCGACA 739
 |||||||
 671 AGTCTGAGGAGACCCGAGAGTGGCCATCATATGTGCGAGAGGAGAAAGAGGCCGACA 730

740 TCTGCGCTCCGAGACAGAAAGGCTGAACAGATTAATTCAGCAGAGAGAGGCGCAGTG 799
 |||||||
 731 TCTGCGCTCCGAGACAGAAAGGCTGAACAGATTAATTCAGCAGAGAGAGGCGCAGTG 790

Db 661 TCCAGGCTACAGATGCAAGT 680

RESULT 7

LOCUS BM461614 1011 bp mRNA linear EST 05-FEB-2002

DEFINITION AGENCOURT_5423435 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5501391

ACCESSION BM461614

VERSION 5', mRNA sequence.

KEYWORDS BM461614.1 GI:18510654

SOURCE EST.

ORGANISM human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 1011)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov

Plate: LLM12137 row: n column: 16

High quality sequence stop: 622.

Location/Qualifiers

1. 1011

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5501391"

/clone_lib="NIH_MGC_67"

/tissue_type="retinoblastoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: oligo Average insert size 1.75 kb. Library constructed by Life Technologies."

Technology: "Life Technologies"

BASE COUNT 232 a 289 c 288 g 201 t 1 others

ORIGIN

Query Match 48 28: Score 573: DB 10: Length 1011:

Best Local Similarity 99.78: Pred. No. 8e-278:

Matches 673: Conservative 0: Mismatches 2: Indels 0: Gaps 0:

Qy 140 GCCCGCCTCTCTGATGGCCGAACACCGTGTCTGTCGCCGAGCAGAGG 199

Db 61 GCCCGCCTCTCTGATGGCCGAACACCGTGTCTGTCGCCGAGCAGAGG 120

Qy 200 CTTGGGTGTGAGCGAATGGCGGATTCACCGGATCCTGGAGCCTGGTTGAACATCC 259

Db 121 CTTGGGTGTGAGCGAATGGCGGATTCACCGGATCCTGGAGCCTGGTTGAACATCC 180

Qy 260 TCATCCCTGTGTAGACCGGATTCGATGTGACAGTCTCAAGGAATGTGCATCAACG 319

Db 181 TCATCCCTGTGTAGACCGGATTCGATGTGACAGTCTCAAGGAATGTGCATCAACG 240

Qy 320 TGCCGTGACAGTGGCTGTGACTCTGACAAATGTAACTCTGCAATGATGGATCTTT 379

Db 241 TGCCGTGACAGTGGCTGTGACTCTGACAAATGTAACTCTGCAATGATGGATCTTT 300

Qy 380 ACCTGCGATCATGAGACCTTACAAAGCAAGTACGGTGTGAGAGACCTTGAGCTGCG 439

Db 301 ACCTGCGATCATGAGACCTTACAAAGCAAGTACGGTGTGAGAGACCTTGAGCTGCG 360

Qy 440 TCACCCAGTACCTCAAAACCAATGATGATGAGTCTGGCAGCAAACTCTCTNTGCAAAAG 499

Db 361 TCACCCAGTACCTCAAAACCAATGATGATGAGTCTGGCAGCAAACTCTCTNTGCAAAAG 420

Qy 500 TCTTCCGGGAACGGAGTCCCTGATGCCAGCATTTGTATGCCATCAACCAAGCTGCTG 559

Db 421 TCTTCCGGGAACGGAGTCCCTGATGCCAGCATTTGTATGCCATCAACCAAGCTGCTG 480

Qy 560 ACTGCTGGGGTATCCGCTGCTTGTATGATGATCAAGATATCCATGTCGCCACCCGGG 619

Db 481 ACTGCTGGGGTATCCGCTGCTTGTATGATGATCAAGATATCCATGTCGCCACCCGGG 540

Qy 620 TGAAGAGTCTATGACATGACAGTGTGAGGACAGCGCGGAAACGGGCGCACATTTCTAG 679

Db 541 TGAAGAGTCTATGACATGACAGTGTGAGGACAGCGCGGAAACGGGCGCACATTTCTAG 600

Qy 680 AGTCTGAGGGGACCCGAGAGTGGCCATCAATGTGCGAAGGAGCAAAACAGGCCGAGA 739

Db 601 AGTCTGAGGGGACCCGAGAGTGGCCATCAATGTGCGAAGGAGCAAAACAGGCCGAGA 660

Qy 740 TCTTGGCCTCCGGAAGCAAGAAAGCTGAACAGATTAATTCAGGACGAGGAGGCGAGTG 799

Db 661 TCTTGGCCTCCGGAAGCAAGAAAGCTGAACAGATTAATTCAGGACGAGGAGGCGAGTG 720

Qy 800 CAGTTCTGGCGAAGG 814

Db 721 CAGTTCTGGCGAAGG 735

RESULT 8

LOCUS BM045426 887 bp mRNA linear EST 07-NOV-2001

DEFINITION 603623541F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:5449259 5', mRNA sequence.

ACCESSION BM045426

VERSION BM045426.1 GI:16774693

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 887)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: DCTD/DTP

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov

Plate: LLM1936 row: b column: 12

High quality sequence stop: 768.

Location/Qualifiers

1. 887

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5449259"

/clone_lib="NIH_MGC_40"

/tissue_type="carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: prostate; Vector: pOT7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)".

Note: this is a NIH_MGC Library."

BASE COUNT 237 a 232 c 247 g 171 t

ORIGIN

Best Local Similarity 99.7%; Pred. No. 1,86-295;
Matches 708; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Oy 203 GGGTGGTGGAGCAATGGCCGATTCACCGGATTCCTGAGACCTGGTTGAACATCTCA 262
    |||||||
Db 1 GGGTGGTGGAGCAATGGCCGATTCACCGGATTCCTGAGACCTGGTTGAACATCTCA 60
Oy 263 TCCCTGTGTAGACCCGGATTCGATATGTGACAGAGTCTCAAGAAATTTGTCATCAACGTGC 322
    |||||||
Db 61 TCCCTGTGTAGACCCGGATTCGATATGTGACAGAGTCTCAAGAAATTTGTCATCAACGTGC 120
Oy 323 CTGAGCAGTCCGCTGTGACCTCTGACATGTACTCTGCAATGCGATGAGTCCCTTACC 382
    |||||||
Db 121 CTGAGCAGTCCGCTGTGACCTCTGACATGTACTCTGCAATGCGATGAGTCCCTTACC 180
Oy 383 TCGCATCATGACCCCTTACAAAGCAAGTACGGTGTGAGAGACCTGAGTATGCGTCA 442
    |||||||
Db 181 TCGCATCATGACCCCTTACAAAGCAAGTACGGTGTGAGAGACCTGAGTATGCGTCA 240
Oy 443 CCCAGCTACCTCAAAACACCATGATGATGAGCTCGGCAAACTCTGTTGACAAAGTCT 502
    |||||||
Db 241 CCCAGCTACCTCAAAACACCATGATGATGAGCTCGGCAAACTCTGTTGACAAAGTCT 300
Oy 503 TCCGGGGAAGGGAGTCCCTGAATGCCAGATTTGGATGCCATCAACCAAGCTGCTGACT 562
    |||||||
Db 301 TCCGGGGAAGGGAGTCCCTGAATGCCAGATTTGGATGCCATCAACCAAGCTGCTGACT 360
Oy 563 GGTGGGATTCGCGTCTGCTTCGTTATGAGATCAAGATATCCATGTGCCACCCGGGTGA 622
    |||||||
Db 361 GGTGGGATTCGCGTCTGCTTCGTTATGAGATCAAGATATCCATGTGCCACCCGGGTGA 420
Oy 623 AAGAGTCTATGACAGATGACAGTGTGAGGACAGCGGCGGAAACGGGCGACAGTTCTAGAGT 682
    |||||||
Db 421 AAGAGTCTATGACAGATGACAGTGTGAGGACAGCGGCGGAAACGGGCGACAGTTCTAGAGT 480
Oy 683 CTGAGGCGACCCGAGAGTGGCCCATCAATCTGGCAGAAAGGGAAGAAACAGGCCCAATCC 742
    |||||||
Db 481 CTGAGGCGACCCGAGAGTGGCCCATCAATCTGGCAGAAAGGGAAGAAACAGGCCCAATCC 540
Oy 743 TGGCCTCCGACAGCAAGAAAGGCTGAACAGATAAATCAGCAGCAGAGAGAGAGCCAGTGCAG 802
    |||||||
Db 541 TGGCCTCCGACAGCAAGAAAGGCTGAACAGATAAATCAGCAGCAGAGAGAGAGCCAGTGCAG 600
Oy 803 TTCTGGCGAAGCGCAAGGCTAAAGCTAATGCTAATTCCTGCTGACAGCTCTGACAG 862
    |||||||
Db 601 TTCTGGCGAAGCGCAAGGCTAAAGCTAATGCTAATTCCTGCTGACAGCTCTGACAG 660
Oy 863 AACATAATGAGATGACAGAGCTTCACTGACTGTGGCCGACAGATATGTC 912
    |||||||
Db 661 AACATAATGAGATGACAGAGCTTCACTGACTGTGGCCGACAGATATGTC 710

```

RESULT 6
BM456748 1053 bp mRNA linear EST 05-FEB-2002
LOCUS BM456748
DEFINITION AGENCOURT_6403923 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5583626
5', mRNA sequence.
ACCESSION BM456748
VERSION BM456748.1 GI:18505788
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 1053)
NIH-MGC <http://mhc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLML12347 row: a column: 03
High quality sequence stop: 621.
Location/Qualifiers

FEATURES
source
1..1053
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5583626"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: pCMV-SPORT6; Site: 1; NotI:
Site: 2; SalI: Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 269 a 285 c 290 g 209 t
ORIGIN

Query Match 48.7%; Score 578; DB 10; Length 1053;
Best Local Similarity 99.7%; Pred. No. 2,46-280;
Matches 678; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Oy 419 TGGAGAACCCCTGGTATGCCGTCACCAGCTAGCTCAAAACCAATGAGATCGACTCG 478
    |||||||
Db 1 TGGAGAACCCCTGGTATGCCGTCACCAGCTAGCTCAAAACCAATGAGATCGACTCG 60
Oy 479 GCAAACTCTCTNTGACAAAGCTTCCGGGAACGGGAGTCCCTGAATGCCAGCATTTGG 538
    |||||||
Db 61 GCAAACTCTCTNTGACAAAGCTTCCGGGAACGGGAGTCCCTGAATGCCAGCATTTGG 120
Oy 539 ATGCCATCAACCAAGCTGCTGACTGCTGGGGTATTCGGCTGCTTCGTTATGAGATCAAG 598
    |||||||
Db 121 ATGCCATCAACCAAGCTGCTGACTGCTGGGGTATTCGGCTGCTTCGTTATGAGATCAAG 180
Oy 599 ATATCATGTGACCCACCCGGGGTAAAGAGTCTATGACAGATGCGAGTGGAGCGCAGACGGC 658
    |||||||
Db 181 ATATCATGTGACCCACCCGGGGTAAAGAGTCTATGACAGATGCGAGTGGAGCGCAGACGGC 240
Oy 659 GGAAGAGGGCCACAGTCTTAAGAGTGAAGGACCCGAGAGTGGCCATCAATGTGGCAG 718
    |||||||
Db 241 GGAAGAGGGCCACAGTCTTAAGAGTGAAGGACCCGAGAGTGGCCATCAATGTGGCAG 300
Oy 719 AAGGGAAGAAACAGGCCAGATCTGGCTCCGAAACGAAAGGCTGAACAGATTAATC 778
    |||||||
Db 301 AAGGGAAGAAACAGGCCAGATCTGGCTCCGAAACGAAAGGCTGAACAGATTAATC 360
Oy 779 AGGCAGCAGAGAGGAGCGAGTTCGAGTCTGGCGAAGGCCAAAGCTTAAGCTTAATTC 838
    |||||||
Db 361 AGGCAGCAGAGAGGAGCGAGTTCGAGTCTGGCGAAGGCCAAAGCTTAAGCTTAATTC 420
Oy 839 GAATCTGCTGCTGAGCTGTGACACAACATTAATGAGATGACAGCAGCTTCACTGACTGTG 898
    |||||||
Db 421 GAATCTGCTGCTGAGCTGTGACACAACATTAATGAGATGACAGCAGCTTCACTGACTGTG 480
Oy 899 CCGAGCAGTATGTACAGCGGCTTCCAAACTGGCCAGGACTCAACACTATCTTACTGC 958
    |||||||
Db 481 CCGAGCAGTATGTACAGCGGCTTCCAAACTGGCCAGGACTCAACACTATCTTACTGC 540
Oy 959 CCTCCAAACCTGGCGATGTACACAGCATGTGGCTGAGCGCATGGGTATATGAGAGGCC 1018
    |||||||
Db 541 CCTCCAAACCTGGCGATGTACACAGCATGTGGCTGAGCGCATGGGTATATGAGAGGCC 600
Oy 1019 TCACCAAGCCCCAGTGCAGAGGACTCCAGACTCTCCAGTGTGGAGCAGCAGAGATG 1078
    |||||||
Db 601 TCACCAAGCCCCAGTGCAGAGGACTCCAGACTCTCCAGTGTGGAGCAGCAGAGATG 660
Oy 1079 TCCAGGGTACAGATGCAAGT 1098
    |||||||

```


BM453159 1061 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT 6387825 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5529785
DEFINITION 5' mRNA sequence.
ACCESSION BM453159
VERSION BM453159.1 GI:18502199
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1061)
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1M12208 row: m column: 18
High quality sequence stop: 661.
Location/Qualifiers
1..1061
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5529785"
/clone_lib="NIH_MGC_71"
/tissue_type="telomysarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: PCMV-SPORT6; Site:1; Noti:
Site:2; Salt: Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb.
BASE COUNT 259 a 298 c 279 g 225 t
ORIGIN
Query Match 54.9%; Score 652; DB 10; Length 1061;
Best Local Similarity 99.7%; Pred. No. 0; Indels 0; Gaps 0;
Matches 752; Conservative 0; Mismatches 2;
OY 345 CGACATGTAAGTTCGCAATGATGAGTCTCTTACTGCGCATGAGCCCTTACAA 404
DB 1 CGACATGTAAGTTCGCAATGATGAGTCTCTTACTGCGCATGAGCCCTTACAA 60
OY 405 GGCAGACTAGCGGTGGAGACCCCTGATGATGCGCTCACCCAGCTCAACACAT 464
DB 61 GGCAGACTAGCGGTGGAGACCCCTGATGATGCGCTCACCCAGCTCAACACAT 120
OY 465 GAGATCAGAGCTCGCAACTCTCTNTGACAAAGCTTCCGGGAGCGGAGTCCCTGAA 524
DB 121 GAGATCAGAGCTCGCAACTCTCTNTGACAAAGCTTCCGGGAGCGGAGTCCCTGAA 180
OY 525 TGCACATGTTGGATGCCATCAACCAAGCTGCTGATGCGGTATCCGCTGCTTCG 584
DB 181 TGCACATGTTGGATGCCATCAACCAAGCTGCTGATGCGGTATCCGCTGCTTCG 240
OY 585 TTATGAGATCAAGATATCCATGTCGCCACCCGGGTGAAGAGTCTATGACATGACGT 644
DB 241 TTATGAGATCAAGATATCCATGTCGCCACCCGGGTGAAGAGTCTATGACATGACGT 300
OY 645 GGAGGAGAGCGGGGGAACCGGGCAGTTCTAGAGTCTGAGGGAGCCGAGAGTCCGC 704
DB 301 GGAGGAGAGCGGGGGAACCGGGCAGTTCTAGAGTCTGAGGGAGCCGAGAGTCCGC 360
OY 705 CATCATGTGGCAGGAAGGAAGCAAGCCAGATCTTGCCGAGCAGAAAAAGGC 764
DB 361 CATCATGTGGCAGGAAGGAAGCAAGCCAGATCTTGCCGAGCAGAAAAAGGC 420

OY 765 TGACAGATTAATTCAGGAGCAGAGAGCCAGTGTCTGCGGAAGCCAGGCTTA 824
DB 421 TGACAGATTAATTCAGGAGCAGAGAGCCAGTGTCTGCGGAAGCCAGGCTTA 480
OY 825 AGCTAAGCTATTCGATTCCTGGCTGAGCTGTGACACAACTAATGAGATGAGCAGC 884
DB 481 AGCTAAGCTATTCGATTCCTGGCTGAGCTGTGACACAACTAATGAGATGAGCAGC 540
OY 885 TTCAGTACTGTGGCGGAGCAGATATGTCAGCGCTTCTCCAACTGAGCCAGACTCCAA 944
DB 541 TTCAGTACTGTGGCGGAGCAGATATGTCAGCGCTTCTCCAACTGAGCCAGACTCCAA 600
OY 945 CACTATCTACTGCGCCCTCCACCCCTGGGATGTCACACAGATGATGCTCAGGCTAGGG 1004
DB 601 CACTATCTACTGCGCCCTCCACCCCTGGGATGTCACACAGATGATGCTCAGGCTAGGG 660
OY 1005 TGTATATGAGGCGCTCACCAGAGCCCGCAGTGGCAGAGACTCCAGACTCAGCTGAG 1064
DB 661 TGTATATGAGGCGCTCACCAGAGCCCGCAGTGGCAGAGACTCCAGACTCAGCTGAG 720
OY 1065 GAGCAGCAGAGATGTCACAGGTTACAGATGCAAGT 1098
DB 721 GAGCAGCAGAGATGTCACAGGTTACAGATGCAAGT 754

RESULT 5
BM423347 1095 bp mRNA linear EST 29-JAN-2002
LOCUS AGENCOURT 6402181 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5516467
DEFINITION 5' mRNA sequence.
ACCESSION BM423347
VERSION BM423347.1 GI:18391559
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1095)
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
TITLE Unpublished (1999)
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTP/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM2017 row: b column: 20
High quality sequence stop: 503.
Location/Qualifiers
1..1095
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5516467"
/clone_lib="NIH_MGC_41"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF7; Site:1; XhoI; Site:2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT 285 a 282 c 302 g 226 t
ORIGIN

Query Match 51.2%; Score 608; DB 10; Length 1095;

FEATURES	source
Db	241 TGAAGTATGCGGTCACCCAGACTAGCTCAACAAACATGAGATAGAGCTGGCAAACTCTC 300
Oy	429 TGAAGTATGCGGTCACCCAGACTAGCTCAACAAACATGAGATAGAGCTGGCAAACTCTC 488
Oy	489 TTTGGCAACAAGTCTTCCGGGAACGGGAGTCCCTGAATGCGACATTTGTGGATCCATCAA 548
Db	301 TCTTGACAAAGCTTCCGGGAACGGGAGTCCCTGAATGCGACATTTGTGGATCCATCAA 360
Oy	549 CCAGACTGCTGACTGCTGGGGTATCCGCTGCCCTNGCTTATGATGATCAAGGATATCATGT 608
Db	361 CCAGACTGCTGACTGCTGGGGTATCCGCTGCCCTTATGATGATCAAGGATATCATGT 420
Oy	609 GCCACCCCGGGTGAAGAAGTCTATGCAGATGAGGTGAGCGAGAGCGGGGAAACGGGC 668
Db	421 GCCACCCCGGGTGAAGAAGTCTATGCAGATGAGGTGAGCGAGAGCGGGGAAACGGGC 480
Oy	669 CACAGTTTCAGAGTCTGAGGGGAGCCGAGAGTCGGCCATCATGTGGGAGGAAGGAAGA 728
Db	481 CACAGTTTCAGAGTCTGAGGGGAGCCGAGAGTCGGCCATCATGTGGGAGGAAGGAAGA 540
Oy	729 ACAGGCCAGATCTGCGCTCCGAAGCAGAAAAAGGCTGAACGATTAATCAGGCGAGCAG 788
Db	541 ACAGGCCAGATCTGCGCTCCGAAGCAGAAAAAGGCTGAACGATTAATCAGGCGAGCAG 600
Oy	789 AGAGGCCAGTGCAGTTCTGGCGCAAGGCCAAGGCTTAAGGTGAAGCTATTTCGAATCCTGGC 848
Db	601 AGAGGCCAGTGCAGTTCTGGCGCAAGGCCAAGGCTTAAGGTGAAGCTATTTCGAATCCTGGC 660
Oy	849 TGCAGCTCTGACACAACATTAATGAGATGACAGCAGCTTCACTGACTGTGGCCGAGCAGTA 908
Db	661 TGCAGCTCTGACACAACATTAATGAGATGACAGCAGCTTCACTGACTGTGGCCGAGCAGTA 720
Oy	909 TGTGAGCGGGTCTCCAAACGTGGCCAGAGACGTCMAAAGCTATCCACACGCTCCCAACCC 968
Db	721 TGTGAGCGGGTCTCCAAACGTGGCCAGAGACGTCMAAAGCTATCCACACGCTCCCAACCC 780
Oy	969 TGGCGATGTCAACCAACATGTTGGCTCAGGCGCATGGGTGTATATGAGAGCCCTCACCAAAAGC 1028
Db	781 TGGCGATGTCAACCAACATGTTGGCTCAGGCGCATGGGTGTATATGAGAGCCCTCACCAAAAGC 840
Oy	1029 CCCAGT 1034
Db	841 CCCAGT 846
RESULT 3	
AL525168	849 bp mRNA linear EST 13-PEB-2001
LOCUS	AL525168 LTI.NFL003.NBC3 Homo sapiens cDNA clone CS0DC005YF20 5
DEFINITION	prime, mRNA sequence.
ACCESSION	AL525168
VERSION	AL525168.1 GI:12788661
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Euteria; Primates; Catarrhini; Hominiidae; Homo.
TITLE	Li, W.-B., Gruber, C., Jesssee, J. and Polayes, D.
JOURNAL	Full-length cDNA libraries and normalization
COMMENT	Unpublished (2001)
	Contact: Genoscope
	Genoscope - Centre National de Sequencage
	BP 191 91006 EVRY cedex - France
	Email: seque@genoscope.cns.fr, web : www.genoscope.cns.fr.
	Location/Qualifiers
	1..849

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DC005YF20"
/clone_lib="LTI_NFL003_NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"

```

Query Match	Best Local Similarity	Score	DB	Length
Matches 760; Conservative	99.7%;	660;	9;	849;
	0;	Mismatches	2;	Indels
	0;	Gaps	0;	
140	GCCTGCTCTCTGATGTTGCCCCGAACACCGTGTACTGTTCTGCTGCCGACGACGAGC	199		
88	GCCTGCTCTCTGATGTTGCCCCGAACACCGTGTACTGTTCTGCTGCCGACGACGAGC	147		
200	CCTGGTGTGAGGAAATGGCGCATTCACCGGATCCTGGAGGCTGGTTGAACATCC	259		
148	CCTGGTGTGAGGAAATGGCGCATTCACCGGATCCTGGAGGCTGGTTGAACATCC	207		
260	TCATCCCTGTGTAGACCGGATCCGATATGTGACAGATCTCAAGGAATTTGCATCAACG	319		
208	TCATCCCTGTGTAGACCGGATCCGATATGTGACAGATCTCAAGGAATTTGCATCAACG	267		
320	TGCCGTGACAGTCCGCTGTGACTCTGCACATGTAACTCTGCAATTCGATGAGTCTTT	379		
268	TGCCGTGACAGTCCGCTGTGACTCTGCACATGTAACTCTGCAATTCGATGAGTCTTT	327		
380	ACCTCGCATATGAGACCCCTTACAAGGCAAGGTAGGTTGAGGACCCCTGATATGCCG	439		
328	ACCTCGCATATGAGACCCCTTACAAGGCAAGGTAGGTTGAGGACCCCTGATATGCCG	387		
440	TCACCCAGCTAGCTCAAAACAACATGAGATCGAGCTCGGCAACTCTCTNTGCAAAAG	499		
388	TCACCCAGCTAGCTCAAAACAACATGAGATCGAGCTCGGCAACTCTCTNTGCAAAAG	447		
500	TCCTTCGGGGAACGGGAGTCCCTGAATGCGCAGCATTTGTGATGTCATCAACCAAGCTGTG	559		
448	TCCTTCGGGGAACGGGAGTCCCTGAATGCGCAGCATTTGTGATGTCATCAACCAAGCTGTG	507		
560	ACTGCTGGGGATATCCGCTGCTTCCTTATGAGATTAAGGATATCCATGTGCCACCCGGG	619		
508	ACTGCTGGGGATATCCGCTGCTTCCTTATGAGATTAAGGATATCCATGTGCCACCCGGG	567		
620	TGAAGAATCTATGAGATGTCAGTGTGAGGCGAGCGCGGAAACGGGCGACAGTTCTAG	679		
568	TGAAGAATCTATGAGATGTCAGTGTGAGGCGAGCGCGGAAACGGGCGACAGTTCTAG	627		
660	AGTCTGAGGGGACCCGAGAGTTCGGCATCAATGTGGCAGAAAGGGAAGAAACAGGCCAGA	739		
628	AGTCTGAGGGGACCCGAGAGTTCGGCATCAATGTGGCAGAAAGGGAAGAAACAGGCCAGA	687		
740	TCCTGCGCTCCGAAACAAAAAGGCTGAAACAGATTAATCAGGACACAGAGGCGCACTG	747		
688	TCCTGCGCTCCGAAACAAAAAGGCTGAAACAGATTAATCAGGACACAGAGGCGCACTG	747		
800	CAGTTCTGGCGAAGGCCAAAGGCTAAAGCTGAAGCTTAATCGAATCCTGGCTGACGCTGGA	859		
748	CAGTTCTGGCGAAGGCCAAAGGCTTAAGGCTGAAGCTTAATCGAATCCTGGCTGACGCTGGA	807		
860	CACAACATTAATGAGATGACAGAGCTTCACTGACTGTGGCG	901		
808	CACAACATTAATGAGATGACAGAGCTTCACTGACTGTGGCG	849		

BASE COUNT

195 a 228 c 260 g 165 t 1 others

ORIGIN

1/lab.host="DB10B"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax: (1) 301 610 8371 Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com"

Query Match

Best Local Similarity

Score

DB

Length

Matches 760; Conservative

99.7%;

660;

9;

849;

0;

Mismatches

2;

Indels

0;

Gaps

0;

140

GCCTGCTCTCTGATGTTGCCCCGAACACCGTGTACTGTTCTGCTGCCGACGACGAGC

199

88

GCCTGCTCTCTGATGTTGCCCCGAACACCGTGTACTGTTCTGCTGCCGACGACGAGC

147

200

CCTGGTGTGAGGAAATGGCGCATTCACCGGATCCTGGAGGCTGGTTGAACATCC

259

148

CCTGGTGTGAGGAAATGGCGCATTCACCGGATCCTGGAGGCTGGTTGAACATCC

207

260

TCATCCCTGTGTAGACCGGATCCGATATGTGACAGATCTCAAGGAATTTGCATCAACG

319

208

TCATCCCTGTGTAGACCGGATCCGATATGTGACAGATCTCAAGGAATTTGCATCAACG

267

320

TGCCGTGACAGTCCGCTGTGACTCTGCACATGTAACTCTGCAATTCGATGAGTCTTT

379

268

TGCCGTGACAGTCCGCTGTGACTCTGCACATGTAACTCTGCAATTCGATGAGTCTTT

327

380

ACCTCGCATATGAGACCCCTTACAAGGCAAGGTAGGTTGAGGACCCCTGATATGCCG

439

328

ACCTCGCATATGAGACCCCTTACAAGGCAAGGTAGGTTGAGGACCCCTGATATGCCG

387

440

TCACCCAGCTAGCTCAAAACAACATGAGATCGAGCTCGGCAACTCTCTNTGCAAAAG

499

388

TCACCCAGCTAGCTCAAAACAACATGAGATCGAGCTCGGCAACTCTCTNTGCAAAAG

447

500

TCCTTCGGGGAACGGGAGTCCCTGAATGCGCAGCATTTGTGATGTCATCAACCAAGCTGTG

559

448

TCCTTCGGGGAACGGGAGTCCCTGAATGCGCAGCATTTGTGATGTCATCAACCAAGCTGTG

507

560

ACTGCTGGGGATATCCGCTGCTTCCTTATGAGATTAAGGATATCCATGTGCCACCCGGG

619

508

ACTGCTGGGGATATCCGCTGCTTCCTTATGAGATTAAGGATATCCATGTGCCACCCGGG

567

620

TGAAGAATCTATGAGATGTCAGTGTGAGGCGAGCGCGGAAACGGGCGACAGTTCTAG

679

568

TGAAGAATCTATGAGATGTCAGTGTGAGGCGAGCGCGGAAACGGGCGACAGTTCTAG

627

660

AGTCTGAGGGGACCCGAGAGTTCGGCATCAATGTGGCAGAAAGGGAAGAAACAGGCCAGA

739

628

AGTCTGAGGGGACCCGAGAGTTCGGCATCAATGTGGCAGAAAGGGAAGAAACAGGCCAGA

687

740

TCCTGCGCTCCGAAACAAAAAGGCTGAAACAGATTAATCAGGACACAGAGGCGCACTG

747

688

TCCTGCGCTCCGAAACAAAAAGGCTGAAACAGATTAATCAGGACACAGAGGCGCACTG

747

800

CAGTTCTGGCGAAGGCCAAAGGCTAAAGCTGAAGCTTAATCGAATCCTGGCTGACGCTGGA

859

748

CAGTTCTGGCGAAGGCCAAAGGCTTAAGGCTGAAGCTTAATCGAATCCT

```

CDS
18. 1046
/codon_start=1
/product="HSPC108"
/protein_id="AF29073.1"
/db_xref="GI:5841440"
/translation="MRSGLASGRACRASSGLPNTVYLVFPOEAMVVERKGRHRI
LEPGLNLIPLVDRIRYVOSLKEIVINVEQSAVTLVDVTLQIDGVLYLRIMDPKAS
YGEDEPEAYTQLAQTMRSELKSLDKYFERESLVAIVDAINQADWQIRCLR
YEKDIHVPKPVESKQMOVAERRRKRTYVESGRTSAINVEAGKKQADILASEAE
KAEIINOAGSAVILAKAKAEIRLILAAILOHNDAAASLTVAROVYASAKLA
KDSNTILPSPNPBDVTSVNAQAMGYGLTRAPVPGTPDLSLSSSSRVOGTDSLDE
ELDRYKMS"

BASE COUNT      313 a      310 c      337 g      258 t
ORIGIN
Query Match      63.6% Score 755; DB 11; Length 1218;
Best Local Similarity 99.6% Pred. No. 0;
Matches 955; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 140 GCCGGCCCTCCCTGGATGCCCCGAAACACCGGTGACTGTTGCGCGCAGCAGAGAG 199
    |||||
DB 52 GCCGGCCCTCCCTGGATGCCCCGAAACACCGGTGACTGTTGCGCGCAGCAGAGAG 111
    |||||

OY 200 CCTGGGTGGAGCGAATGGCCGATTCACCGGATCCTGGAGCCTGGTTGAACATCC 259
    |||||
DB 112 CCTGGGTGGAGCGAATGGCCGATTCACCGGATCCTGGAGCCTGGTTGAACATCC 171
    |||||

OY 260 TCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAAATGTTCATCAAG 319
    |||||
DB 172 TCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAAATGTTCATCAAG 231
    |||||

OY 320 TCCGTGAGCAGTGGGTGTGATCTGTGACATGTACTGTGCAAAATGATGAGTCTT 379
    |||||
DB 232 TCCGTGAGCAGTGGGTGTGATCTGTGACATGTACTGTGCAAAATGATGAGTCTT 291
    |||||

OY 380 ACCGTGGCATCATGACACCTTACAGAGCAAGCTACGGTGTGGAGACCTGAGTATGCCG 439
    |||||
DB 292 ACCGTGGCATCATGACACCTTACAGAGCAAGCTACGGTGTGGAGACCTGAGTATGCCG 351
    |||||

OY 440 TCACCAGCTAGCTCAAAACCATGATGATGAGTCCGCAAACTCTGNTGACAAAG 499
    |||||
DB 352 TCACCAGCTAGCTCAAAACCATGATGATGAGTCCGCAAACTCTCTGACAAAG 411
    |||||

OY 500 TCTTCCGGAGAGGAGTCCCTGTAATGCCAGATTGTGATGCCATCAACCAAGTCTG 559
    |||||
DB 412 TCTTCCGGAGAGGAGTCCCTGTAATGCCAGATTGTGATGCCATCAACCAAGTCTG 471
    |||||

OY 560 ACTGCTGGGGTATCCGCTGCTTCGTTATGAGATCAAGGATATTCATGTCACCCGGG 619
    |||||
DB 472 ACTGCTGGGGTATCCGCTGCTTCGTTATGAGATCAAGGATATTCATGTCACCCGGG 531
    |||||

OY 620 TGAAGAGTCTATGAGATGACAGTGGAGGAGCGCGGAAAGGGGACAGTTCTAG 679
    |||||
DB 532 TGAAGAGTCTATGAGATGACAGTGGAGGAGCGCGGAAAGGGGACAGTTCTAG 591
    |||||

OY 680 AGTCTGAGGGAGCCGAGAGTCCGATCAATGTGGCAGAGGAAAGAAACAGGCCAGA 739
    |||||
DB 592 AGTCTGAGGGAGCCGAGAGTCCGATCAATGTGGCAGAGGAAAGAAACAGGCCAGA 651
    |||||

OY 740 TCTGTGGCTCCGAGAGGAAAGGCTGAACAGATTAATTCAGGAGGAGAGGCCAGTG 799
    |||||
DB 652 TCTGTGGCTCCGAGAGGAAAGGCTGAACAGATTAATTCAGGAGGAGAGGCCAGTG 711
    |||||

OY 800 CATTTCGGCGAAGCGCAAGGCTAAAGCTATTCGAATCTTGAGTCCAGCTCTGA 859
    |||||
DB 712 CATTTCGGCGAAGCGCAAGGCTAAAGCTATTCGAATCTTGAGTCCAGCTCTGA 771
    |||||

OY 860 CACAACATAATGAGATGACAGAGCTTCACATGACTGSCCGAGCAAGTATGTCAGCCGT 919
    |||||
DB 772 CACAACATAATGAGATGACAGAGCTTCACATGACTGSCCGAGCAAGTATGTCAGCCGT 831
    |||||

OY 920 TCTCCAAACTGGCGAAGAGCTCCACACACTATCTACTGCCCTCCAAACCTGGCGATGTCA 979
    |||||

```

```

DB 832 TCTCCAAACTGGCGAAGAGCTCCACACTATCTCTGCTCCCAACCTGGCGATGTCA 891
    |||||
OY 980 CCAGCATGTGCTCAGGCGCATGAGTGTATGAGAGCCCTCACCAAGGCCAGTCCAG 1039
    |||||
DB 892 CCAGCATGTGCTCAGGCGCATGAGTGTATGAGAGCCCTCACCAAGGCCAGTCCAG 951
    |||||

OY 1040 GGACTCCAGACTCTCTCCAGTGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGT 1098
    |||||
DB 952 GGACTCCAGACTCTCTCCAGTGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGT 1010
    |||||

RESULT 2
AL541792      935 bp mRNA linear EST 16-FEB-2001
LOCUS      AL541792 LTI_FL002_PL1 Homo sapiens cDNA clone CS0DE007ye18 5 prime
DEFINITION      mRNA sequence.
ACCESSION      AL541792.1 GI:12873201
VERSION      AL541792.1
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 935)
AUTHORS      Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT      Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 9191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr; Web : www.genoscope.cns.fr.

FEATURES
source
1..935
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="LTI_FL002_PL1"
/lab_host="DH10B"
/note="Organ: Placenta; Vector: pCMVSPORT 6; 1st strand
cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-stranded cDNA was digested with Not I
and cloned into the Not I and Eco RV sites of the
pCMVSPORT 6 vector. Library was constructed by life
technologies. Contact : Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
, Maryland 20850, USA Fax : (1) 301 610 8371 Email :
liang@life.com URL :
http://fulllength.invitrogen.com"

BASE COUNT      239 a      246 c      265 g      181 t      4 others
ORIGIN
Query Match      58.3% Score 693; DB 9; Length 935;
Best Local Similarity 99.6% Pred. No. 0;
Matches 843; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 189 GCAGCAGAGGCTGGGTGTGAGAGCGAATGGCCGATTCACCCGATCTCGAGCTCG 248
    |||||
DB 1 GCAGCAGAGGCTGGGTGTGAGAGCGAATGGCCGATTCACCCGATCTCGAGCTCG 60
    |||||

OY 249 TTTGAACATCCTCATCCCTGTTAGACGGGATCCGATGTGTCAGAGTCTCAAGGAAT 308
    |||||
DB 61 TTTGAACATCCTCATCCCTGTTAGACGGGATCCGATGTGTCAGAGTCTCAAGGAAT 120
    |||||

OY 309 TGTCAATCAAGTGTGAGAGAGTGGCTGTGACTCTTCGACAAATGTAACTGCAAAATGCA 368
    |||||
DB 121 TGTCAATCAAGTGTGAGAGAGTGGCTGTGACTCTTCGACAAATGTAACTGCAAAATGCA 180
    |||||

OY 369 TGGAGTCTTTTACCTGGCGATCATGAGACCTTACAAAGCAAGCTACGCTGTGAGAGACC 428
    |||||
DB 181 TGGAGTCTTTTACCTGGCGATCATGAGACCTTACAAAGCAAGCTACGCTGTGAGAGACC 240
    |||||

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 15:22:53 ; Search time 1701.18 Seconds
(without alignments)
9425.447 Million cell updates/sec

Title: US-09-898-216-2

Perfect score: 1188
Sequence: 1 GCGTTCGAGACNACCCT.....GGACGAGATTTCCTGATT 1188

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: em_estda:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pin:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	755	63.6	1218	11	AF161458 Homo sapi
2	693	58.3	935	9	AL541792 AL541792
3	660	55.6	849	9	AL525168 AL525168
4	652	54.9	1061	10	BM453159 BM453159
5	608	51.2	1095	10	BM423347 BM423347
6	578	48.7	1053	10	BM456748 BM456748
7	573	48.2	1011	10	BM461614 BM461614
8	572	48.1	887	10	BM454526 BM454526
9	571	48.1	1110	10	BM476304 BM476304
10	568	47.8	796	10	BM1092184 BM1092184
11	558	47.0	660	10	BE740453 BE740453
12	555	46.7	702	10	BI838402 BI838402
13	552	46.5	749	10	BC036463 BC036463
14	539	45.4	770	10	BC828070 BC828070
15	536	45.1	1037	10	BM471641 BM471641
16	534	44.9	853	10	BE314702 BE314702
17	531	44.7	820	10	BG685336 BG685336

18	518	43.6	994	10	BM468527 BM468527
19	516	43.4	918	10	BM459766 BM459766
20	515	43.4	644	10	BE273795 BE273795
21	513	43.2	616	10	BM045444 BM045444
22	504	42.4	614	10	BE256589 BE256589
23	504	42.4	688	10	BE684435 BE684435
24	501	42.2	728	10	BI522493 BI522493
25	500	42.1	689	10	BM015578 BM015578
26	500	42.1	920	10	BI862105 BI862105
27	498	41.9	962	10	BG112719 BG112719
28	487	41.0	998	10	BM468407 BM468407
29	483	40.7	737	10	BG700808 BG700808
30	483	40.7	841	10	BE904561 BE904561
31	482	40.6	739	10	BG574270 BG574270
32	478	40.2	934	10	BF344401 BF344401
33	477	39.7	730	10	BI761597 BI761597
34	472	39.7	738	10	BG767007 BG767007
35	471	39.6	778	10	BG681190 BG681190
36	470	39.6	752	10	BI552983 BI552983
37	469	39.5	785	10	BM009759 BM009759
38	464	39.1	836	10	BE734242 BE734242
39	463	39.0	691	10	BE297150 BE297150
40	463	39.0	693	10	BF311329 BF311329
41	463	39.0	707	10	BI533283 BI533283
42	463	39.0	725	10	BE275386 BE275386
43	463	39.0	736	10	BI198047 BI198047
44	463	39.0	845	10	BE893277 BE893277
45	463	39.0	871	10	BE794548 BE794548

ALIGNMENTS

RESULT 1	AF161458	1218 bp	MRNA	linear	HTC 22-MAY-2001
LOCUS	AF161458				
DEFINITION	Homo sapiens HSPC108 mRNA, complete cds.				
ACCESSION	AF161458				
VERSION	AF161458.1	GI:6841439			
KEYWORDS	HTC.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Zhang,Q.H., Ye,M., Wu,X.Y., Ren,S.X., Zhao,M., Zhao,C.J., Fu,G., Shen,Y., Fan,H.Y., Lu,G., Zhong,M., Xu,X.R., Han,Z.G., Zhang,J.W., Tao,J., Huang,Q.H., Zhou,J., Hu,G.X., Gu,J., Chen,S.J., and Chen,Z.				
TITLE	Cloning and functional analysis of cDNAs with open reading frames for 300 previously undefined genes expressed in CD34+ hematopoietic stem/progenitor cells				
JOURNAL	Genome Res. 10 (10), 1546-1560 (2000)				
MEDLINE	20499367				
PUBMED	11042152				
REFERENCE	2 (bases 1 to 1218)				
AUTHORS	Ye,M., Zhang,Q.H., Zhou,J., Shen,Y., Wu,X.Y., Guan,Z.Q., Wang,L., Fan,H.Y., Mao,Y.F., Dai,M., Huang,Q.H., Chen,S.J., and Chen,Z.				
TITLE	Human full length cDNA cloned from cd34+ stem cells				
JOURNAL	Unpublished				
REFERENCE	3 (bases 1 to 1218)				
AUTHORS	Ye,M., Zhang,Q.H., Zhou,J., Shen,Y., Wu,X.Y., Guan,Z.Q., Wang,L., Fan,H.Y., Mao,Y.F., Dai,M., Huang,Q.H., Chen,S.J., and Chen,Z.				
TITLE	Direct Submission				
JOURNAL	Submitted (23-JUN-1999) Shanghai Institute of Hematology, Shanghai Second Medical University, Rui-Jin Hospital, 197 Rui-Jin Road II, Shanghai 200025, P. R. China				
FEATURES	location/Qualifiers				
source	1..1218				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="CBCAPB12"				
	/cell_type="cd34+ stem cells"				
	/tissue_type="blood"				

PRIOR FILLING DATE: 1997-06-13
 PRIOR APPLICATION NUMBER: 60/049,609
 PRIOR FILLING DATE: 1997-06-13
 PRIOR APPLICATION NUMBER: 60/049,610
 PRIOR FILLING DATE: 1997-06-13
 PRIOR APPLICATION NUMBER: 60/049,611
 PRIOR FILLING DATE: 1997-06-13
 PRIOR APPLICATION NUMBER: 60/050,901
 PRIOR FILLING DATE: 1997-06-13
 PRIOR APPLICATION NUMBER: 60/052,989
 PRIOR FILLING DATE: 1997-06-13
 PRIOR APPLICATION NUMBER: 60/051,919
 PRIOR FILLING DATE: 1997-07-08
 PRIOR APPLICATION NUMBER: 60/055,984
 PRIOR FILLING DATE: 1997-08-18
 PRIOR APPLICATION NUMBER: 60/058,665
 PRIOR FILLING DATE: 1997-09-12
 PRIOR APPLICATION NUMBER: 60/058,668
 PRIOR FILLING DATE: 1997-09-12
 PRIOR APPLICATION NUMBER: 60/058,669
 PRIOR FILLING DATE: 1997-09-12
 PRIOR APPLICATION NUMBER: 60/058,750
 PRIOR FILLING DATE: 1997-09-12
 PRIOR APPLICATION NUMBER: 60/058,971
 PRIOR FILLING DATE: 1997-09-12
 PRIOR APPLICATION NUMBER: 60/058,972
 PRIOR FILLING DATE: 1997-09-12
 PRIOR APPLICATION NUMBER: 60/058,975
 PRIOR FILLING DATE: 1997-09-12
 PRIOR APPLICATION NUMBER: 60/060,834
 PRIOR FILLING DATE: 1997-10-02
 PRIOR APPLICATION NUMBER: 60/060,841
 PRIOR FILLING DATE: 1997-10-02
 PRIOR APPLICATION NUMBER: 60/060,844
 PRIOR FILLING DATE: 1997-10-02
 PRIOR APPLICATION NUMBER: 60/060,865
 PRIOR FILLING DATE: 1997-10-02
 PRIOR APPLICATION NUMBER: 60/061,059
 PRIOR FILLING DATE: 1997-10-02
 PRIOR APPLICATION NUMBER: 60/061,060
 PRIOR FILLING DATE: 1997-10-02
 NUMBER OF SEQ ID NOS: 737
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 106
 LENGTH: 1322
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (82)
 OTHER INFORMATION: n equals a,t,g, or c
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (1320)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (1321)
 OTHER INFORMATION: n equals a,t,g, or c
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (1322)
 OTHER INFORMATION: n equals a,t,g, or c
 US-10-219-793-106

Query Match 72.0%; Score 855; DB 6; Length 1322;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 955; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Db 142 CGGCGCTCTCTGATGCGCCGAAACACCGTGTACTGTTGCGCGACAGAGAGGCC 201
 139 CGGCGCTCTCTGATGCGCCGAAACACCGTGTACTGTTGCGCGACAGAGAGGCC 198

QY 202 TGGTGTGAGCAATGGGCCGATTCACCGGATTCCTGAGAGCCTGTTGAACATCTC 261
 Db 199 TGGTGTGAGCAATGGGCCGATTCACCGGATTCCTGAGAGCCTGTTGAACATCTC 258
 QY 262 ATCCCTGTTAGACCGGATTCGATGTCAGAGTCTCAAGAAATGTCATCAACGTG 321
 Db 259 ATCCCTGTTAGACCGGATTCGATGTCAGAGTCTCAAGAAATGTCATCAACGTG 318
 QY 322 CCTGAGCAGTGTGCTGATCTGACATGTCAGATGTCAGATGTCAGATGTCAGATGTC 381
 Db 319 CCTGAGCAGTGTGCTGATCTGACATGTCAGATGTCAGATGTCAGATGTCAGATGTC 378
 QY 382 CTGGGCATCATGACCCCTTAAGGCAAGCTACGCTGTGAGAGACCCCTGATGCGCTG 441
 Db 379 CTGGGCATCATGACCCCTTAAGGCAAGCTACGCTGTGAGAGACCCCTGATGCGCTG 438
 QY 442 ACCGAGCTAGCTCAACCAACCATGATGATGATGATGATGATGATGATGATGATGATG 501
 Db 439 ACCGAGCTAGCTCAACCAACCATGATGATGATGATGATGATGATGATGATGATGATG 498
 QY 502 TTCCGGAGACGGAGTCCCTGATGATGATGATGATGATGATGATGATGATGATGATG 561
 Db 499 TTCCGGAGACGGAGTCCCTGATGATGATGATGATGATGATGATGATGATGATGATG 558
 QY 562 TGTGAGGAGACCGGAGTCCCTGATGATGATGATGATGATGATGATGATGATGATGATG 621
 Db 559 TGTGAGGAGACCGGAGTCCCTGATGATGATGATGATGATGATGATGATGATGATGATG 618
 QY 622 AAAGAGTCTATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 681
 Db 619 AAAGAGTCTATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 678
 QY 682 TGTGAGGAGACCGGAGTCCCTGATGATGATGATGATGATGATGATGATGATGATGATG 741
 Db 679 TGTGAGGAGACCGGAGTCCCTGATGATGATGATGATGATGATGATGATGATGATGATG 738
 QY 742 CTGGGCTCCGAGCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 801
 Db 739 CTGGGCTCCGAGCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 798
 QY 802 GTTGTGCGAAGCCCAAGGCTAAAGCTGAAGCTATGATGATGATGATGATGATGATG 861
 Db 799 GTTGTGCGAAGCCCAAGGCTAAAGCTGAAGCTATGATGATGATGATGATGATGATG 858
 QY 862 CAACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 921
 Db 859 CAACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 918
 QY 922 TCCCAAGTGGCCAGGATGCAACACTATGCTAGTCCCTCCCAAGCTGGAGTGCACC 981
 Db 919 TCCCAAGTGGCCAGGATGCAACACTATGCTAGTCCCTCCCAAGCTGGAGTGCACC 978
 QY 982 AGCATGCTGCTACGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1041
 Db 979 AGCATGCTGCTACGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1038
 QY 1042 ACTCCAGACTACTCTCCAGTGGAGACAGACAGATGTCAGAGTCAAGTCAAGTCAAGT 1098
 Db 1039 ACTCCAGACTACTCTCCAGTGGAGACAGATGTCAGAGTCAAGTCAAGTCAAGT 1095

RESULT 7
 US-09-629-469A-11009
 ; Sequence 11009, Application US/09629469A
 ; GENERAL INFORMATION:
 ; APPLICANT: OTA, TOSHIO
 ; APPLICANT: ISOGAI, TAKAO
 ; APPLICANT: NISHIKAWA, TETSUO
 ; APPLICANT: HAYASHI, KOJI
 ; APPLICANT: SAITO, KAORI
 ; APPLICANT: YAMAMOTO, JUNICHI
 ; APPLICANT: ISHII, SHIZUKO

Db 920 tctcaaacctgagcccaaggactccaacactctcactgcccctccaacccctgagatgtca 979
Qy 980 CCAAGCATGTGTGGCTCAGGCCATGGGTATATGAGAGCCCTCACCAGAAAGCCCAAGTGGCAG 1039
Db 980 ccaagcatgtgtgctcagagccatgggtgtatataagagccctcccaaaagccccagtgccag 1039
Qy 1040 GGAAGCTCAGACTCAGCTTCCAGTGGGAGCAGAGAGATGTCCAGGGTACAGATGCAAGT 1098
Db 1040 ggaactcagaactcactctccagtgtagagcagcagaagatgctcagggtacagatgtcaagt 1098

RESULT 5

US-10-119-428-37
; Sequence 37, Application US/10119428
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Wehrman, Tom
; APPLICANT: Ren, Felyan
; APPLICANT: Ma, Yungling
; APPLICANT: Zhou, Ping
; APPLICANT: Zhao, Qiang A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: 789CIP2
; CURRENT APPLICATION NUMBER: US/10/119,428
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PL_FL_genes Version 1.0
; SEQ ID NO 37
; LENGTH: 9098
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (75)..(9098)
US-10-119-428-37

Query Match 72.18; Score 857; DB 7; Length 9098;
Best Local Similarity 99.88; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 140 GCGGCGCTCTCTGATGATCCCGGAAACACCGGTACTGTCGTCGCCGACGAGGAGG 199
Db 151 gccggccctctcttgatctgcccgaacacgcgtgtaactctcgctgcgcagaagaagg 210
Qy 200 CCTGGGTGTGAGCAGATGGGCCGATTCACCGGATCTGAGCCCTGGTTGAACATCC 259
Db 211 cctgggtgtgtagcagatgggccgatctcacccggtatcccgagcccggtttgaactcc 270
Qy 260 TCATCCCTGTGTAGACCGGATCCGATATGTGACAGACTCTCAAGAAATTGTATCAACG 319
Db 271 tcatccctgtgttagaccggaatccgatatgtcagaagttccaagaaattgtcatcaag 330
Qy 320 TGGCTGAGAGCGGCTGTGACTCTGACAAATGTAATCTGCAATGTGATGGAGTCTCTT 379
Db 331 tggctgagagcgtgtgactctgcacaaatgtaactctgcaaatgtatgaggtccctt 390
Qy 380 ACCTGGCATCATGAGCCCTTACAGCAAGCTAAGGCTGTGAGAGACCTGATATGATCCG 439
Db 391 acctggcatcatagacccttacaagcaagctaaggtgtgtgagagaccctgataagcgg 450

Qy 440 TCACCCAGCTAGCTCAAAACACATGAGATCAGAGCTCGGCAAACTCTNTGCAACAAG 499
Db 451 tcaaccagctagctcaaaacacatgagatcagagcttcgccaactctctctgtgaaag 510
Qy 500 TCTTCCGGGAACGGGACTCCCTTGAATGCCAGCATTTGTGATGTCCATTAACCAAGCTCTG 559
Db 511 tcttcgggaacggaagctcccttgaaatgcagcatgtgtgatatgcaacaacagctgtcg 570
Qy 560 ACTGCTGGGTATCCGCTGCTTCGCTTATGATGATCAAGGATATTCATGTGCCACCCGGG 619
Db 571 actgctgggtatccgcttcgcttcctcgttatgatatcaagatatcatgtgtcaaccggg 630
Qy 620 TGAAGAGCTTATGACATCAGTCAGTGAGCGCAGACGGCGAAAGCGCCACAGTTCTAG 679
Db 631 tgaagagcttatgtcagatgcaggtgtaggcagagcgcggaacggtccacagttctag 690
Qy 680 AGTCTGAGGGAGCCCGAGCTCGGCTCATGANTGTGGCAGAAAGGAACAAACAGCCCGAGA 739
Db 691 agtctgagggagcccgagctcggtccatcaatgtgcaagaaggaaagaaacagccaga 750
Qy 740 TCTGAGCTCCGAAGCAGAAAGGCTGACAGATTAATTCAGGCGAGAGAGGCCAGTG 799
Db 751 tctgagctccgaagcagaaagctgtacaagataaattcagcagcagagagggccagtg 810
Qy 800 CAGTTCTGGCGAAGCCCAAGGCTTAAGCTTAAGCTTATTCATCTGCTGCAAGCTCTGA 859
Db 811 cagttctggcgaagcccaaggtcaagctgaagctatctcgaatctctgtgcagctctga 870
Qy 860 CACACATTAATGGAGATGAGAGAGCTTCACTGACTGTGGCGAGCATATGTACGGCGT 919
Db 871 cacaacataatgagatgagcagcttcacagctgtgcccagcagatgataagcgcgt 930
Qy 920 TCTCCAACTGGCCCAAGAGACTCCACACTATCTACTGCCCCCTCAACCTGCGGATGTCA 979
Db 931 tctccaactggccaaggaactccaactatctctacgtccctccaaacctcgagatgtca 990
Qy 980 CCAAGCATGTGTGCTCAGGCCATGGGTATATGAGAGCCCTCACCAGAAAGCCCAAGTGGCAG 1039
Db 991 ccaagcatgtgtgctcagagccatgggtgtatataagagccctcacaagagcccgatgcccag 1050
Qy 1040 GGAAGCTCAGACTCAGCTTCCAGTGGGAGCAGAGAGATGTCCAGGGTACAGATGCAAGT 1098
Db 1051 ggaactcagaactcactctccagtgtagagcagcagaagatgctcagggtacagatgtcaagt 1109

RESULT 6

US-10-219-793-106
; Sequence 106, Application US/10219793
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; FILE REFERENCE: P2008P1C1
; CURRENT APPLICATION NUMBER: US/10/219,793
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/209,462
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: PCT/US98/12125
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/049,547
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,548
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,549
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,550
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,566
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,606
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,607
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,608

Db 679 cctgggtggtgagcgaatggtggtccacggatctccgtgagcctgttgaaatcc 738
 QY 260 TCATCCTGTGTAGACCGGATCCGATATGTGACAGTCTCAAGAAATGTGCATCAACG 319
 Db 739 tcatccctgtgttaacccggatccgatattgtcagaagctcgaagaattgtcacaag 798
 QY 320 TGCCTGACAGCTGGCTGTGACTCTGACAAATGTAACTCTGCAAAATCGATGAGTCTTT 379
 Db 799 tgcctgagcagctgagctgtgactctcgaacaatgtaactctgaaatcgatgagctt 858
 QY 380 ACCCTGGATCATGACCTTTCAGAGCAAGTACGAGTGTGAGAGACCTGTGATGCCG 439
 Db 859 accctgcatctgtgaccccttaagaagctacagctgtgagagccctgtgctgccc 918
 QY 440 TCACCAGCTACCTCAAAACCATGAGATGAGCTGAGCTGGCAACTCTCTNNAGCAAAAG 499
 Db 919 tcaaccagctagctcaaaacacatgagatcagagctggcaaatctctcgtgacaag 978
 QY 500 TCTTCGGGAGACGGAGTCCCTGAATGCCAGCATTGTGGATGCCATCAACCAAGCTGTG 559
 Db 979 tcttcgggaacggagctccctgaaatgccaagctgtgagtgcatcaaccagctgtg 1038
 QY 560 ACTGCTGGGATTCGGCTGCTTNCCTTATGATCAAGATTCATGTGCCACCCGGG 619
 Db 1039 actgctgggtatccgctgctcgttaltgagatcaagataatccatgtgccacccgg 1098
 QY 620 TGAAGAGCTATGAGATGAGTGAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 679
 Db 1099 tgaagagctctatgagatgagctgagagcagagcagagcagagcagagcagagct 1158
 QY 680 ACTGTGAGGAGCCGAGAGTGGCCATCATATGTGGCAGAAAGGAGAAACAGGCCGAGA 739
 Db 1159 agctgagaggagccgagagctgcgcacatcatgtgagcagaaggagaaacagccaga 1218
 QY 740 TCCCTGGCTCCCAAGCAGAAAGGCTGAACAGATTAATAGCAGCAGAGAGAGGCGCAG 799
 Db 1219 tccctggctccgaagcagaaagctgaaacagataaatacagacagagagagcagag 1278
 QY 800 CAGTCTGGGAGGAGGAGGCTAAAGCTGAAGCTATTGGAATCCGCTGAGAGCTTGA 859
 Db 1279 cagttctggcgaagcgaagctaaagctgaagctatctgaaatccctgctgacgtctga 1338
 QY 860 CACAACATATGAGATGACAGAGCTTCACTGACTGTGGCCGAGCATATGTCAGCGGT 919
 Db 1339 cacaacataatgagatgagcagagctcaatgagctgtgagcagagatgacagcgct 1398
 QY 920 TCTCCAACTGGCCAGGAGCTCCAACTATCTACTGCTCCCAACCTGGCGATGTCA 979
 Db 1399 tctccaaactggcgaagagctccaaacatactcctactgcccctccaaacccctggagatgca 1458
 QY 980 CCAGCATGGTGGCTCAGGCGATGGGTGTATGAGAGCCCTCACCAAGGCCAGTGCAG 1039
 Db 1459 ccagcatggtgctcagcagcagtgatataatgagcctccaccaaagccacagtgca 1518
 QY 1040 GCACTTCACACTCTCTCCAGTGGGAGCAGCAGACATGTCCAGGTTACAGATGCAAGT 1098
 Db 1519 gactccacagctactcctcagtgagagcagagagatgtccagggtaacagatgcaagt 1577

US-10-198-846-12975
 ; Sequence 12975, Application US/10198846
 ; GENERAL INFORMATION:
 ; APPLICANT: Lillie, James
 ; APPLICANT: Xu, Yongyao
 ; APPLICANT: Wang, Youzhen
 ; APPLICANT: Steinmann, Kathleen
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
 ; TITLE OF INVENTION: FOR IDENTIFICATION OF BREAST CANCER
 ; FILE REFERENCE: MRI-049
 ; CURRENT APPLICATION NUMBER: US/10/198,846
 ; CURRENT FILING DATE: 2002-07-18

RESULT 4

US-10-198-846-12975

Query Match 72.1%; Score 857; DB 6; Length 1910;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 140 GCCGCGCTCTCTGTGATGTCGCCGAAACCGGTGATCTGTGCTGCCGAGCAGAGG 199
 Db 140 gccgcgctctctgtgatgtgctccgaaacacggtgactgtgtgctgcccagcagag 199
 QY 200 CCTGGGTGAGGAGGAGATGGCGCATTCACCGGATCTGAGGCTGGTTGAACATCC 259
 Db 200 cctgggtgagagcagatggcgcatctcagagcagatccagagctgtgttgaaatcc 259
 QY 260 TCATCCTGTGTAGACCGGATCCGATATGTGAGAGTCAAGGAAATGTGCATCAACG 319
 Db 260 tcatccctgtgttagaccgagatccgatattgtcagaagctcgaagaattgtcacaag 319
 QY 320 TGCCTGACAGCTGGCTGTGACTCTGACAAATGTAACTCTGCAAAATCGATGAGTCTTT 379
 Db 320 tgcctgacagctggctgtgactctcgaacaatgtaactctgaaatcgatgagctt 379
 QY 380 ACCCTGGATCATGACCTTTCAGAGCAAGTACGAGTGTGAGAGACCTGTGATGCCG 439
 Db 380 accctggatcatgaccttcaagcagatcagcgtgtgagagacccctgtgagctgag 439
 QY 440 TCACCAGCTACCTCAAAACCATGAGATGAGCTGAGCTGGCAACTCTCTNNAGCAAAAG 499
 Db 440 tcaaccagctagctcaaaacacatgagatcagagctggcaaatctctcgtgacaag 499
 QY 500 TCTTCGGGAGACGGAGTCCCTGAATGCCAGCATTGTGGATGCCATCAACCAAGCTGTG 559
 Db 500 tcttcgggagacggagtccctgaaatgccaagctgtgagtgcatcaaccagctgtg 559
 QY 560 ACTGCTGGGATTCGGCTGCTTNCCTTATGATCAAGATTCATGTGCCACCCGGG 619
 Db 560 actgctgggatctcggctgcttnccttattgagatcaagatcacaatgtgccaacccgg 619
 QY 620 TGAAGAGCTATGAGATGAGTGAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 679
 Db 620 tgaagagctatgagatgagtgagagctgagagcagagcagagcagagcagagctgag 679
 QY 680 AGTCTGAGGAGCAGGAGAGTGGCCATTAATGTGCGAGAGGAGAAACAGAGCCCA 739
 Db 680 agtctgagggagccagagagtgcccatatattgtgcagaaaggaaacagggccaga 739
 QY 740 TCTTCGGCTCCGAGCAGAAAGGCTGAACAGTGAATCAAGCAGCAGCAGCAGCAGT 799
 Db 740 tcttcgctccgagcagaaaggctgaacagtgaacagataacagcagcagagagccagtg 799
 QY 800 CAGTCTGGCAGAGGCGCTAAAGCTGAACCTATTGCAATCTGCTGCAAGCTCTGA 859
 Db 800 cagttctggcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcga 859
 QY 860 CACAACATATGAGATGACAGAGCTTCACTGACTGTGGCCGAGCATATGTCAGCGGT 919
 Db 860 cacaacataatgagatgagcagagctcaatgagctgtgagcagagatgtgtaagcgct 919
 QY 920 TCTCCAACTGGCCAGGAGCTCCAACTATCTACTGCTCCCAACCTGGCGATGTCA 979
 Db 920 tctccaaactggcgaagagctccaaactatctactgctcccaacccctggagatgca 979

```

; PRIOR APPLICATION NUMBER: 60/060,844
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,865
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,059
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,060
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 737
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 45
; LENGTH: 1337
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1335)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1336)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1337)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-219-793-45
```

Query Match 72.1% Score 857; DB 6; Length 1337;

Best Local Similarity 99.8%; Pred. No. 0; Mismatches 2; Indels 0; Gaps 0;

```

Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 140 GCCGCCCTCTCTGTGATTGCCCGAACAACCGTGTACTGTTGCGCCAGCAGAGG 199
    |||||
DB 152 gccgcgcccctctgtgattgccccgaacaacggtgtactgttcgycgcagcaggaag 211
QY 200 CCTGGTGTGAGCGAATGGCCGATTCCACCGGATCCTGAGCCTGTTGAACATCC 239
    |||||
DB 212 ccgaggctgtgagcgaaatgagcgatccacgcgattccctgagccctgttgaaatcc 271
QY 260 TCATCCCTGTGTAGACCGGATTCGATNTGTGAGAGTCTCAGAGAAATTTGTATCAAG 319
    |||||
DB 272 tcatccctgtgttagaccggaatccgatalgtgagagctccaaggaattgtatcaacg 331
QY 320 TCCCTGAGAGTGTGCGTGTGACTCTGACAAATGTAATCTGCAAAATGATGAGTCTTT 379
    |||||
DB 332 tgcctgagcagctcgctgtgacctctgacaatgttaactctgcaaatcgatgggtcctt 391
QY 380 ACCTGGCATCATGAGACCTTACAGGCAAGCTAGCGTGTGAGAGAACCTGAGTATGCCG 439
    |||||
DB 392 acctgcatcatgagacccttaacaagcaagctacggtgtgaggaagccctgagctatg 451
QY 440 TCACCGAGTAGTCAAAACCATGATGATCAAGCTCGGCAAACTCTGNTGCAAAAG 499
    |||||
DB 452 tcacccaagctcgaacaacaacgaagatcgaagctcggcaaacctctcttgaaacaaag 511
QY 500 TCTTCCGAGACGAGTCCCTGAATGCGCAGCATTTGTGATGCCATCAACCAAGTCTGTG 559
    |||||
DB 512 tcttcgggaaagggagctccctgaatgcccagctgtgtgagctgcatcaacaagcgtgtg 571
QY 560 ACTGCTGGGGTATCCGCTGCTTCGTTATGAGATCAAGGATATCATGTGCCACCCGGGG 619
    |||||
DB 572 actgctgggggtatccgctcgtctcgtctatgagatcaaggtatcctatgtgccaccgg 631
QY 620 TCAAGAGCTATGAGAGATGAGAGTGGAGGCAAGCGCGGAAACGGGCCACAGTTCTAG 679
    |||||
DB 632 tgaagaagctctatgcaagatgcaagggtgagggcagagcgagaaacgggcccacgttctag 691
QY 680 ACTCTGAGGAGACCCGAGAGTCCGCCATGATGTGGCAGAGAGGAGAAACAGGCCGAGA 739
    |||||
DB 692 agtctgagggagaccggagagctgggccaatgtgtgcaagaaggaaacagggcccaaga 751
```

```

QY 740 TCCTGGCTCCGAGCAGAGAAAGCTGMAACGATTAATTCAGGACGAGAGGCCAGTG 799
    |||||
DB 752 tcttgccctcgaagcagaaaggctgaaacagataaatacgaagcagagagccaggtg 811
QY 800 CATTTCGGCGAAGGCCAAGGCTTAAGCTTAAGCTATTGCAATTCCTGGCTGCAGCTCTGA 859
    |||||
DB 812 cagltctggaggaagccaaagctaaagctgaagctatcgaatcttcggctgcagcttga 871
QY 860 CACAACATATGAGATGTCAGCAGCTTCATGACTGTGCGCGGAGCAGATATGTCAGCGGT 919
    |||||
DB 872 cacaacaataatgagatgacgcagcttcaactgactgtgtgcgcagctatgtaagcgct 931
QY 920 TCTCCAAACTGGCCCAAGACTCCACACTATCTACTGCCCTCCAAACCTGGCGATGTCA 979
    |||||
DB 932 tctccaactgccaagactccaacactatcttactgtccctccaacacctgycgagtca 991
QY 980 CCAGCATGTGTGCTCAGAGCCATGGGTATATATGAGAGCCCTCACCMAAGCCCAAGTGCAG 1039
    |||||
DB 992 ccagcatgtgtctcagcgcaatgtgtatagagccctcaccacaaagcccgctccag 1051
QY 1040 GGAATCCAGACTCACTCTCCAGTGGGAGCAGAGATGTCAGAGGATGATGCAAGT 1098
    |||||
DB 1052 ggaatccagactcaactcctccagctggtgagcagagatgtccaggttaccagatgcaagt 1110
```

RESULT 3

US-09-785-276A-24973

; Sequence 24973, Application US/09785276A

; GENERAL INFORMATION:

; APPLICANT: Schlegel, Robert

; APPLICANT: Endege, Wilson

; APPLICANT: Monahan, John

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR

; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF

; TITLE OF INVENTION: HUMAN PROSTATE CANCER

; FILE REFERENCE: MRI-007B

; CURRENT APPLICATION NUMBER: US/09/785,276A

; PRIOR FILING DATE: 2001-02-16

; PRIOR APPLICATION NUMBER: 60/183,319

; PRIOR FILING DATE: 2000-02-17

; PRIOR APPLICATION NUMBER: 60/189,862

; PRIOR FILING DATE: 2000-03-16

; PRIOR APPLICATION NUMBER: 60/207,454

; PRIOR FILING DATE: 2000-05-25

; PRIOR APPLICATION NUMBER: 60/211,314

; PRIOR FILING DATE: 2000-06-09

; PRIOR APPLICATION NUMBER: 60/219,007

; PRIOR FILING DATE: 2000-07-18

; PRIOR APPLICATION NUMBER: 60/255,281

; PRIOR FILING DATE: 2000-12-13

; NUMBER OF SEQ ID NOS: 62232

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 24973

; LENGTH: 1842

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: 1841..1842

; OTHER INFORMATION: n = A,T,C or G

US-09-785-276A-24973

Query Match 72.1% Score 857; DB 5; Length 1842;

Best Local Similarity 99.8%; Pred. No. 0; Mismatches 2; Indels 0; Gaps 0;

Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 140 GCCGCCCTCTCTGTGATTGCCCGAACAACCGTGTACTGTTGCGCCAGCAGAGG 199
    |||||
DB 619 gccgcgcccctctgtgattgccccgaacaacggtgtactgttcgycgcagcaggaag 678
QY 200 CCTGGTGTGAGCGAATGGCCGATTCCACCGGATCTGAGAGCTGTTGAACATCC 259
    |||||
```

[illegible]

```

Db      1021  ACCAAGGCCCAAGGCCAGGAGACTCCAGACTCACTCCAGGTGGAGACGACGAGATATTC 10800
QY      1081  CAGGCTACAGATGCAGAGTNTTGATGTAGAGACTTGATCGATCGATCAAGATGAGTACTGAGAC 11400
Db      1081  CAGGCTACAGATGCAGAGTNTTGATGTAGAGACTTGATCGATCGATCAAGATGAGTACTGAGAC 11400
QY      1141  TGGGCTTGGCCAGGAGGTCTGGGGACACAGACACATTTTCTGATT 1188
Db      1141  TGGGCTTGGCCAGGAGGTCTGGGGACACAGACATTTTCTGATT 1188

RESULT 2
US-10-219-793-45
: Sequence 45, Application US/10219793
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: 86 Human Secreted Proteins
: FILE REFERENCE: P2008P1C1
: CURRENT APPLICATION NUMBER: US/10/219,793
: CURRENT FILING DATE: 2002-08-16
: PRIOR APPLICATION NUMBER: 09/209,462
: PRIOR FILING DATE: 1998-12-11
: PRIOR APPLICATION NUMBER: PCT/US98/12125
: PRIOR FILING DATE: 1998-06-11
: PRIOR APPLICATION NUMBER: 60/049,547
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,548
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,549
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,550
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,566
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,606
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,607
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,608
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,609
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,610
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,611
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/050,901
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/052,989
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/051,919
: PRIOR FILING DATE: 1997-07-08
: PRIOR APPLICATION NUMBER: 60/055,984
: PRIOR FILING DATE: 1997-08-18
: PRIOR APPLICATION NUMBER: 60/058,665
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,668
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,669
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,750
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,971
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,972
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,975
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/060,834
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/060,841
: PRIOR FILING DATE: 1997-10-02

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 16:48:58 : Search time 322.14 Seconds
(without alignments)
11832.429 Million cell updates/sec

Title: US-09-898-216-2

Perfect score: 1188
Sequence: 1 GGCTTCTGGAGACNACCGCT.....GGAAGACAGATTTCCTGATT 1188

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 2138461 seqs, 1604250230 residues

Word size: 0

Total number of hits satisfying chosen parameters: 4276922

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: Pending Patents_NA_New.*

- 1: /cgn2_6/ptodata/2/pna/US06_NEM_COMB.seq.*
- 2: /cgn2_6/ptodata/2/pna/US06_NEM_COMB.seq.*
- 3: /cgn2_6/ptodata/2/pna/US07_NEM_COMB.seq.*
- 4: /cgn2_6/ptodata/2/pna/US08_NEM_COMB.seq.*
- 5: /cgn2_6/ptodata/2/pna/US09_NEM_COMB.seq.*
- 6: /cgn2_6/ptodata/2/pna/US10_NEM_COMB.seq.*
- 7: /cgn2_6/ptodata/2/pna/US10_NEM_COMB.seq.*
- 8: /cgn2_6/ptodata/2/pna/US10_NEM_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1188	100.0	1188	US-09-898-216-2	Sequence 2, Appl
2	857	72.1	1337	US-10-219-793-45	Sequence 45, Appl
3	857	72.1	1842	US-09-785-276A-24973	Sequence 24973, A
4	857	72.1	1910	US-10-198-846-12975	Sequence 12975, A
5	857	72.1	9098	US-10-119-428-37	Sequence 37, Appl
6	855	72.0	1322	US-10-219-793-106	Sequence 106, Appl
7	806	67.8	1244	US-09-629-468A-11009	Sequence 11009, A
8	674	56.7	1234	US-09-703-256A-3468	Sequence 3468, Ap
9	616	51.9	1150	US-10-198-846-13409	Sequence 13409, A
C 10	572	48.1	1309	US-09-919-002-11863	Sequence 11863, A
C 11	509	42.8	622	PCT-US01-43704-1161	Sequence 1161, Ap
C 12	454	38.2	454	US-10-146-502-193	Sequence 193, Ap
C 13	453	38.1	550	US-10-214-403-2895	Sequence 2895, Ap
C 14	402	33.8	504	PCT-US01-43704-841	Sequence 841, Appl
C 15	391	32.9	563	PCT-US01-43704-58	Sequence 58, Appl
C 16	390	32.8	448	US-09-920-300A-1075	Sequence 1075, Ap
C 17	390	32.8	448	US-10-099-926-1075	Sequence 1075, Ap
C 18	357	30.1	357	US-10-214-403-2873	Sequence 2873, Ap
C 19	352	29.6	439	PCT-US02-25766-1798	Sequence 1798, Ap
C 20	331	27.9	691	US-10-097-105-1428	Sequence 1428, Ap
C 21	316	26.6	445	US-09-918-995-5335	Sequence 5335, Ap
C 22	279	23.5	362	US-09-721-544-8507	Sequence 8507, Ap
C 23	267	22.5	478	US-09-918-995-31316	Sequence 31316, A
C 24	262	22.1	430	US-09-785-276A-37723	Sequence 37723, A
C 25	253	21.3	235	US-09-539-331D-4825	Sequence 4825, Ap

C 26	250	21.0	301	1	PCT-US02-12378-26	Sequence 26, Appl
C 27	250	21.0	301	7	US-10-124-805-26	Sequence 26, Appl
C 28	221	18.6	221	5	US-09-539-331D-15714	Sequence 15714, A
C 29	221	18.6	221	5	US-09-539-331D-21791	Sequence 21791, A
C 30	216	18.2	253	5	US-09-539-331D-688	Sequence 688, Appl
C 31	211	17.8	211	5	US-09-540-210B-18465	Sequence 18465, A
C 32	199	16.8	199	5	US-09-540-210B-6996	Sequence 6996, Ap
C 33	171	14.4	567	5	US-09-629-468A-1950	Sequence 1950, Ap
C 34	169	14.2	740	7	US-10-097-105-596	Sequence 596, Appl
C 35	160	13.5	239	5	US-09-539-331D-16278	Sequence 16278, A
C 36	151	12.7	232	5	US-09-539-331D-12906	Sequence 12906, A
C 37	149	12.5	300	5	US-09-539-331D-9558	Sequence 9558, A
C 38	148	12.5	227	5	US-09-539-331D-15486	Sequence 15486, A
C 39	147	12.4	282	5	US-09-540-210B-7087	Sequence 7087, Ap
C 40	144	12.1	214	5	US-09-540-210B-15126	Sequence 15126, A
C 41	139	11.7	265	5	US-09-539-331D-14079	Sequence 14079, A
C 42	125	10.5	390	5	US-09-785-276A-7770	Sequence 7770, Ap
C 43	124	10.4	902	6	US-10-198-846-6734	Sequence 6734, Ap
C 44	122	10.3	242	5	US-09-539-331D-30091	Sequence 30091, A
C 45	122	10.3	1158	6	US-10-198-846-5820	Sequence 5820, Ap

ALIGNMENTS

RESULT 1
US-09-898-216-2
Sequence 2, Application US/09898216
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
Goll, Surya K.
TITLE OF INVENTION: NOVEL HUMAN MEMBRANE PROTEIN
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/898, 216
FILING DATE: 02-Jul-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/781,562
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0181 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <unknown>
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1188 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE: Consensus
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-898-216-2


```

: PRIOR FILLING DATE: 1999-01-29
: PRIOR APPLICATION NUMBER: US 09/248,797
: PRIOR FILLING DATE: 1999-02-12
: PRIOR APPLICATION NUMBER: US 09/271,490
: PRIOR FILLING DATE: 1999-03-18
: PRIOR APPLICATION NUMBER: US 09/293,972
: PRIOR FILLING DATE: 1999-04-15
: PRIOR APPLICATION NUMBER: US 09/274,861
: PRIOR FILLING DATE: 1999-03-23
: PRIOR APPLICATION NUMBER: US 60/125,453
: PRIOR FILLING DATE: 1999-03-19
: PRIOR APPLICATION NUMBER: US 60/126,605
: PRIOR FILLING DATE: 1999-03-26
: PRIOR APPLICATION NUMBER: US 09/306,350
: PRIOR FILLING DATE: 1999-05-07
: PRIOR APPLICATION NUMBER: US 09/399,720
: PRIOR FILLING DATE: 1999-09-21
: PRIOR APPLICATION NUMBER: US 09/404,284
: PRIOR FILLING DATE: 1999-09-21
: PRIOR APPLICATION NUMBER: US 09/465,877
: PRIOR FILLING DATE: 1999-12-16
: PRIOR APPLICATION NUMBER: NOT YET ASSIGNED
: PRIOR FILLING DATE: 1999-01-19
: NUMBER OF SEQ ID NOS: 10410
: SOFTWARE: PC_CT_genes Version 1.02
: SEQ ID NO: 8410
: LENGTH: 1398
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1296)..(264)
: OTHER INFORMATION: similar to g12984585 in the genepept database release 114,
US-09-496-914A-8410

```

```

Query Match          72.1%; Score 857; DB 18; Length 1398;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 140 GCGGGCCCTCCCTGCGATGCGCCCAACACCGGTGCTGCTGCGCGGACGAGAG 199
Db 1254 GCGGGCCCTCCCTGCGATGCGCCCAACACCGGTGCTGCTGCGCGGACGAGAG 1195

Qy 200 CCGGGGTGGAGCGAATGGGCGCATTCACCGGATCCTGGAGCCTGTTGAACATCC 259
Db 1194 CCGGGGTGGAGCGAATGGGCGCATTCACCGGATCCTGGAGCCTGTTGAACATCC 1135

Qy 260 TCATCCCTGTGTAGACCGGATCGATATGTGACAGAGTCTCAAGAAATGTCAATCAACG 319
Db 1134 TCATCCCTGTGTAGACCGGATCGATATGTGACAGAGTCTCAAGAAATGTCAATCAACG 1075

Qy 320 TGGCTGACAGCGGCGTGTGACCTGCAATGTAATGCAATGCAATGCAATGCAATGCAATG 379
Db 1074 TGGCTGACAGCGGCGTGTGACCTGCAATGTAATGCAATGCAATGCAATGCAATGCAATG 1015

Qy 380 ACCTCGCAATCATGAGACCTTACAGGCAAGCTAGCTGTGAGAGACCTGAGTATGCCG 439
Db 1014 ACCTCGCAATCATGAGACCTTACAGGCAAGCTAGCTGTGAGAGACCTGAGTATGCCG 955

Qy 440 TCACCCAGCTAGCTCAAAACACCATGAGATCAGAGCTGGCAAATCTCTTGTGACAAG 499
Db 954 TCACCCAGCTAGCTCAAAACACCATGAGATCAGAGCTGGCAAATCTCTTGTGACAAG 895

Qy 500 TCTTCGCGGAGGAGGAGTCCGTAATGCGAGATGCGAGATGCGAGATGCGAGATGCGAGATG 559
Db 894 TCTTCGCGGAGGAGGAGTCCGTAATGCGAGATGCGAGATGCGAGATGCGAGATGCGAGATG 835

Qy 560 ACTGCTGGGGTATCGCGTCCCTGCTATGAGATCAAGGATATCCATGTGCCACCCGGG 619
Db 834 ACTGCTGGGGTATCGCGTCCCTGCTATGAGATCAAGGATATCCATGTGCCACCCGGG 775

```

```

Qy 620 TGAAGACTCTATGAGATGAGTGGAGGAGCGGCGGAAACGGGCGACAGTTCTAG 679
Db 774 TGAAGACTCTATGAGATGAGTGGAGGAGCGGCGGAAACGGGCGACAGTTCTAG 715

Qy 680 AGCTGAGGGGACCCGAGAGTGGCGCATCATGTGGCAGAAAGGAGAAACAGGCCACGA 739
Db 714 AGCTGAGGGGACCCGAGAGTGGCGCATCATGTGGCAGAAAGGAGAAACAGGCCACGA 655

Qy 740 TCTGCGCTCCGAGACAGAAAGGCTGAACAGATTAATTCAGGACAGAGAGGCGCACTG 799
Db 654 TCTGCGCTCCGAGACAGAAAGGCTGAACAGATTAATTCAGGACAGAGAGGCGCACTG 595

Qy 800 CAGTTCTGGCGAAGCGCAAGGCTAAAGCTGAAGCTATTGCAATCTGGCTGACGCTGGA 859
Db 594 CAGTTCTGGCGAAGCGCAAGGCTAAAGCTGAAGCTATTGCAATCTGGCTGACGCTGGA 535

Qy 860 CACAACATTAATGAGATGACAGAGCTTACAGTGTGTGGCGGAGAGATATGTCAAGCGGT 919
Db 534 CACAACATTAATGAGATGACAGAGCTTACAGTGTGTGGCGGAGAGATATGTCAAGCGGT 475

Qy 920 TCTCCAACTGGCCCAAGGACTCCAACTATCTACTGCTCCAACTGCGGATGTCA 979
Db 474 TCTCCAACTGGCCCAAGGACTCCAACTATCTACTGCTCCAACTGCGGATGTCA 415

Qy 980 CCAGCATGGTGGCTCAGGCCATGGGTGTATATGAGCCCTCACCAGCCCGAGTCCAG 1039
Db 414 CCAGCATGGTGGCTCAGGCCATGGGTGTATATGAGCCCTCACCAGCCCGAGTCCAG 355

Qy 1040 GAGCTCCAGACTCACTCTCCAGTGGGAGACAGAGATGTCCAGGGTACAGATGCAAT 1098
Db 354 GAGCTCCAGACTCACTCTCCAGTGGGAGACAGAGATGTCCAGGGTACAGATGCAAMGT 296

```

Search completed: September 22, 2002, 18:14:09
Job time: 6820 sec

```

; EARLIER APPLICATION NUMBER: US 09/234,611
; EARLIER FILING DATE: 1999-01-22
; EARLIER APPLICATION NUMBER: US 09/240,371
; EARLIER FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: US 09/277,227
; EARLIER FILING DATE: 1999-03-25
; EARLIER APPLICATION NUMBER: US 09/271,490
; EARLIER FILING DATE: 1999-03-18
; EARLIER APPLICATION NUMBER: US 09/293,972
; EARLIER FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: US 09/274,861
; EARLIER FILING DATE: 1999-03-23
; EARLIER APPLICATION NUMBER: US 60/125,453
; EARLIER FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: US 60/126,605
; EARLIER FILING DATE: 1999-03-26
; EARLIER APPLICATION NUMBER: US 09/306,350
; EARLIER FILING DATE: 1999-05-07
; EARLIER APPLICATION NUMBER: US 09/399,720
; EARLIER FILING DATE: 1999-09-21
; EARLIER APPLICATION NUMBER: US 09/404,284
; EARLIER FILING DATE: 1999-09-21
; EARLIER APPLICATION NUMBER:
; NUMBER OF SEQ ID NOS: 10451
; SOFTWARE: PC-CT_genes Version 1.0
; SEQ ID NO: 865
; LENGTH: 1398
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1296)...(264)
; OTHER INFORMATION: similar to g12984585 in the genepept database release 114,
; OTHER INFORMATION: Run with FASTX 3.3t00, default parameters
US-09-471-275-865
```

```

Query Match          72.1%; Score 857; DB 18; Length 1398;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 140 GCCGCGCCCTCTCTGATGATGCCCCGAACACCGGTGCTGTTCTGCGCAGCAGAGG 199
    |||||
DB 1254 GCCGCGCCCTCTCTGATGATGCCCCGAACACCGGTGCTGTTCTGCGCAGCAGAGG 1195
    |||||
QY 200 CTTGGGTGTGAGGAATGGGCCCATTCACCGGATCTCTGAGACCTGGTTGAACATCC 259
    |||||
DB 1194 CTTGGGTGTGAGGAATGGGCCCATTCACCGGATCTCTGAGACCTGGTTGAACATCC 1135
    |||||
QY 260 TCATCCCTGTGTAGACCGGATCGATATGTGCAGAGTCTCAAGGAATTTGTCATCAACG 319
    |||||
DB 1134 TCATCCCTGTGTAGACCGGATCGATATGTGCAGAGTCTCAAGGAATTTGTCATCAACG 1075
    |||||
QY 320 TGCCTGAGCAGTGGCTGTGACTCTCGACAATGTAACTCTGCAAAATCGATGAGTCTTT 379
    |||||
DB 1074 TGCCTGAGCAGTGGCTGTGACTCTCGACAATGTAACTCTGCAAAATCGATGAGTCTTT 1015
    |||||
QY 380 ACCTGCGCATATGAGACCTTACAAGGCAAGCTAGGGTGGAGGACCTGAGTATGCCG 439
    |||||
DB 1014 ACCTGCGCATATGAGACCTTACAAGGCAAGCTAGGGTGGAGGACCTGAGTATGCCG 955
    |||||
QY 440 TCACCGAGCTACTCAAAACCAATGAGATGAGTCTGCGCAAACTCTTNNGCAAAAG 499
    |||||
DB 954 TCACCGAGCTACTCAAAACCAATGAGATGAGTCTGCGCAAACTCTTNNGCAAAAG 895
    |||||
QY 500 TCTTCCGGGAAGGGAGTCCCTGAATGCCAGCATTTGTGATGCCATCAACCAAGCTGCTG 559
    |||||
DB 894 TCTTCCGGGAAGGGAGTCCCTGAATGCCAGCATTTGTGATGCCATCAACCAAGCTGCTG 835
    |||||
QY 560 ACTGCTGGGGTATCCGCTGCTTNCSTTATGAGATCAAGGATATTCATGTGCCACCCGGG 619
    |||||
DB 834 ACTGCTGGGGTATCCGCTGCTTNCSTTATGAGATCAAGGATATTCATGTGCCACCCGGG 775
    |||||
```

```

QY 620 TGAAGAGTCTATGACATGACAGTGTGAGGACAGAGCGGGAACGGGCCACAGTTCTAG 679
    |||||
DB 774 TGAAGAGTCTATGACATGACAGTGTGAGGACAGAGCGGGAACGGGCCACAGTTCTAG 715
    |||||
QY 680 AGTGTGAGGGGACCCGAGAGTGGCCCATCATATGTGGCAGAGAAGGAAGAAAGGCCCA 739
    |||||
DB 714 AGTGTGAGGGGACCCGAGAGTGGCCCATCATATGTGGCAGAGAAGGAAGAAAGGCCCA 655
    |||||
QY 740 TCTTGCCCTCCCAAGCAGAAAGGCTGAACACATTAATCAGCAGCAGAGAGGCCACAG 799
    |||||
DB 654 TCTTGCCCTCCCAAGCAGAAAGGCTGAACACATTAATCAGCAGCAGAGAGGCCACAG 595
    |||||
QY 800 CAGTTCTGCGGAAGCCCAAGGCTAAAGCTGAAGCTATTGCAATCCGGCTGCAGCTCTGA 859
    |||||
DB 594 CAGTTCTGCGGAAGCCCAAGGCTAAAGCTGAAGCTATTGCAATCCGGCTGCAGCTCTGA 535
    |||||
QY 860 CACAACATATATGAGATGAGCAGAGCTTCACTGACTGTGGCCGAGCAGATATGTCAGCGCT 919
    |||||
DB 534 CACAACATATATGAGATGAGCAGAGCTTCACTGACTGTGGCCGAGCAGATATGTCAGCGCT 475
    |||||
QY 920 TCTCCAACTGGCCCAAGGACTCCACACATATCTGCTCCCAACCTGGCGATGTGA 979
    |||||
DB 474 TCTCCAACTGGCCCAAGGACTCCACACATATCTGCTCCCAACCTGGCGATGTGA 415
    |||||
QY 980 CCAGCATGTGGCTCAGGCCATGGGTGTATATGAGACCCCTCACCAAGGCCAGTGCAG 1039
    |||||
DB 414 CCAGCATGTGGCTCAGGCCATGGGTGTATATGAGACCCCTCACCAAGGCCAGTGCAG 355
    |||||
QY 1040 GGACTTCAGACTCTCTCCAGTGGGAGCAGCAGACAGATGTCCAGGGTACAGATGCAAGT 1098
    |||||
DB 354 GGACTTCAGACTCTCTCCAGTGGGAGCAGCAGACAGATGTCCAGGGTACAGATGCAAGT 296
    |||||
```

RESULT 15

```

US-09-496-914A-8410/C
; Sequence 8410, Application US/09496914A
; GENERAL INFORMATION:
; APPLICANT: Tang, Yuanhua T.
; APPLICANT: Tillinghast, John
; APPLICANT: Sinku, Ankura
; APPLICANT: Liu, Chenghua
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Conlfigs Obtained
; FILE REFERENCE: From Various Libraries
; FILE REFERENCE: 787
; CURRENT APPLICATION NUMBER: US/09/496,914A
; CURRENT FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/353,690
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/034,341
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 09/045,400
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 09/321,214
; PRIOR FILING DATE: 1999-05-26
; PRIOR APPLICATION NUMBER: US 09/131,598
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: US 09/431,517
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: US 09/328,351
; PRIOR FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: US 09/332,782
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/235,076
; PRIOR FILING DATE: 1999-01-20
; PRIOR APPLICATION NUMBER: US 09/234,611
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: US 09/346,956
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: US 09/362,510
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: US 09/240,371
```

QY 680 AGCTGAGGGACCCGAGAGTCGCCATCATGTGGCGAGAAAGGAGAAACAGGCCCCAGA 739
DB 714 AGTCTGAGGGGACCCGAGAGTCGCCATCATGTGGCGAGAAAGGAGAAACAGGCCCCAGA 655
QY 740 TCCGTGGCTCCGAGACGACAAAAGGCTGAACAGATTAATCAGGACACAGAGAGGCGCTAG 799
DB 654 TCTGTGGCTCCGAGACGACAAAAGGCTGAACAGATTAATCAGGACACAGAGAGGCGCTAG 595
QY 800 CAGTTCTGGCGAAGGCGCAAGGCTAAAGCTGAAGCTATTGCAATCTGGCTGACGCTCTGA 859
DB 594 CAGTTCTGGCGAAGGCGCAAGGCTAAAGCTGAAGCTATTGCAATCTGGCTGACGCTCTGA 535
QY 860 CACAACATTAATGAGATGACAGCAGCTTCACTGACTGTGGCGAGCAGATATGTCAAGCGGT 919
DB 534 CACAACATTAATGAGATGACAGCAGCTTCACTGACTGTGGCGAGCAGATATGTCAAGCGGT 475
QY 920 TCTCCAACTGGCGCAAGGACCTCCACACTATCTGCTCCCAACCTGGCGATGTCA 979
DB 474 TCTCCAACTGGCGCAAGGACCTCCACACTATCTGCTCCCAACCTGGCGATGTCA 415
QY 980 CCAGCATGTGGCTCAGGCGCATGGGTATATGAGAGCCCTCACCAAGCCCGAGTCCAG 1039
DB 414 CCAGCATGTGGCTCAGGCGCATGGGTATATGAGAGCCCTCACCAAGCCCGAGTCCAG 355
QY 1040 GGACTCCAGACTCTCTCCACTGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGT 1098
DB 354 GGACTCCAGACTCTCTCCACTGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGT 296

RESULT 13
PCT-US01-04098A-2212/c
; Sequence 2212, Application PC/TUS0104098A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: PCT/US01/04098A
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: Not yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 2212
; LENGTH: 1398
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US01-04098A-2212

Query Match 72.1%; Score 857; DB 1; Length 1398;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 140 GCCGCCCTCTCTGTGATTGCCCGAAGACCGTGTACTGTCTGTCGCGCAGCAGAGAG 199
DB 1254 GCCGCCCTCTCTGTGATTGCCCGAAGACCGTGTACTGTCTGTCGCGCAGCAGAGAG 1195

QY 200 CCTGGGTGGTGGAGGAATGGGCGCGATTCCACCGGATTCCTTGAGCTGTTGAACATCC 259
DB 1194 CTTGGTGGTGGAGGAATGGGCGCGATTCCACCGGATTCCTTGAGCTGTTGAACATCC 1135
QY 260 TCATCCCTGTGTAGACCGGATTCGATATGTGAGAGTCTCAAGAAATTTCTATCAACG 319
DB 1134 TCATCCCTGTGTAGACCGGATTCGATATGTGAGAGTCTCAAGAAATTTCTATCAACG 1075
QY 320 TGCCAGACAGTGGCTGTGACTCTGACAAATGTAACTGTGCAATGTGAGATGCTTT 379
DB 1074 TGCCAGACAGTGGCTGTGACTCTGACAAATGTAACTGTGCAATGTGAGATGCTTT 1015
QY 380 ACCTGCGATCATGAGACCTTACAAAGCAGCTAGCGTGTGAGAGACCTGAGTATGCG 439
DB 1014 ACCTGCGATCATGAGACCTTACAAAGCAGCTAGCGTGTGAGAGACCTGAGTATGCG 955
QY 440 TCACCAGCTAGCTCAAAACACATGAGATCAGAGCTGGCAACTCTCTGTGACAAAG 499
DB 954 TCACCAGCTAGCTCAAAACACATGAGATCAGAGCTGGCAAACTCTCTGTGACAAAG 895
QY 500 TCTTCGGGGAAGGGAGTCCCTGAATGCGAGCATTTGTGATGCCATCAACCAAGCTCTG 559
DB 894 TCTTCGGGGAAGGGAGTCCCTGAATGCGAGCATTTGTGATGCCATCAACCAAGCTCTG 835
QY 560 ACTGCTGGGATATCCGCTGCTTCGTTATGAGATCAAGATATCCATGTGCCACCCCGG 619
DB 834 ACTGCTGGGATATCCGCTGCTTCGTTATGAGATCAAGATATCCATGTGCCACCCCGG 775
QY 620 TGAAGAGCTATGAGATGAGTGAAGCAGAGCGGCGGAAAGGGGCGACAGTTCTAG 679
DB 774 TGAAGAGCTATGAGATGAGTGAAGCAGAGCGGCGGAAAGGGGCGACAGTTCTAG 715
QY 680 AGCTGAGGGGACCCGAGAGTGGCCATCATGTGCGACAGAGGAAAGAAACAGGCCCCAGA 739
DB 714 AGCTGAGGGGACCCGAGAGTGGCCATCATGTGCGACAGAGGAAAGAAACAGGCCCCAGA 655
QY 740 TCTGTGGCTCCGAGACGACAAAAGGCTGAACAGATTAATCAGGACACAGAGAGGCGCTAG 799
DB 654 TCTGTGGCTCCGAGACGACAAAAGGCTGAACAGATTAATCAGGACACAGAGAGGCGCTAG 595
QY 800 CAGTTCTGGCGAAGGCGCAAGGCTAAAGCTGAAGCTATTTGAAATCTGCTGCGAGTCTGA 859
DB 534 CACAACATTAATGAGATGACAGCAGCTTCACTGACTGTGGCGAGCAGATATGTCAAGCGGT 475
QY 920 TCTCCAACTGGCGCAAGGACCTCCACACTATCTGCTCCCAACCTGGCGATGTCA 979
DB 474 TCTCCAACTGGCGCAAGGACCTCCACACTATCTGCTCCCAACCTGGCGATGTCA 415
QY 980 CCAGCATGTGGCTCAGGCGCATGGGTATATGAGAGCCCTCACCAAGCCCGAGTCCAG 1039
DB 414 CCAGCATGTGGCTCAGGCGCATGGGTATATGAGAGCCCTCACCAAGCCCGAGTCCAG 355
QY 1040 GGACTCCAGACTCTCTCCACTGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGT 1098
DB 354 GGACTCCAGACTCTCTCCACTGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGT 296

RESULT 14
US-09-471-275-865/c
; Sequence 865, Application US/09471275
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Contigs Obtained
; FILE REFERENCE: 782
; CURRENT APPLICATION NUMBER: US/09/471,275
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 09/235,076
; EARLIER FILING DATE: 1999-01-20

	Query Match	72.1%	Score 857	DB 25	Length 1384
	Best Local Similarity	99.8%	Pred. No. 0		
	Matches 957	Conservative	0	Indels 2	Gaps 0
QY	140	GGCGGCGCTCTGTGATTGCCCCGAAACCGTGTACTGTTTCGTCGGCGGACGACAGAGG	199		

```

RESULT 11
US-09-808-384-941
; Sequence 941, Application US/09808384
; GENERAL INFORMATION:
; APPLICANT: Gearling, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN PROSTATE STROMAL LIBRARY
; FILE REFERENCE: 1600.1025-002
; CURRENT APPLICATION NUMBER: US/09/808,384

```

|||||
Db 918 tcccaaacgycgaaggactccaacactcctactcgcctcccaacccgycgaltca 977
QY 980 CCAGCATGTGGCTCAGGCCATGGGTATATGAGCCCTCACCAGAGCCCGAGTGCAG 1039
Db 978 ccagcatggtggtctcagccatggtgtatataagagccctccacaagccccagtgccag 1037
QY 1040 GGACTCCAGACTACTCTCCAGTGGAGCAGCAGAGATGTCACAGGATACAGATGCAAGT 1098
Db 1038 ggaactccagactcactctccagctggtgagcagcagagatgtccaggtgtacagatgcaagt 1096

RESULT 8

US-09-652-816-7419
; Sequence 7419, Application US/09652816
; GENERAL INFORMATION:
; APPLICANT: Gutierrez-Ramos, Jose-Carlos
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1177-001
; CURRENT APPLICATION NUMBER: US/09/652,816
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,111
; NUMBER OF SEQ ID NOS: 9647
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 7419
; LENGTH: 1384
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-816-7419

Query Match 72.1%; Score 857; DB 25; Length 1384;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 140 GCCGCGCTCTCTGTGATTGCCCCGAAACACCGTGTACTGTTCTGTGCCGACAGAGAG 199
Db 138 gccgcgcctcctctgtatgctcccgaaacacgctgtactgtctgcgcgacagagag 197
QY 200 CCGGGTGTGGAGCCAAATGGCCGATTCACCGGATCCTGGAGCCCTGGTTGAACATCC 259
Db 198 cctcggtgtggtgagcgaatggtgacgtatccacggatccctggagccctgttgaacatcc 257
QY 260 TCATCCCTGTGTAGACCGGATCGATATGTGAGAGTCTCAAGAAATGTGATCAACG 319
Db 258 tcctccctgtgttagacggatccgatatgtgcagagctcccaagaaatgtgatacaag 317
QY 320 TGCCTGACAGTGGCTGTGACTCTCGACAAATGAATGATGAGTCCCTTT 379
Db 318 tgcctgagcagctgctgtgactctcgacaatgttaactctgcaaatcgatgagctctt 377
QY 380 ACCTGGCATGAGGACCCCTTACAAGCAAGTACGATGTGGAGGACCCCTGATATGCCG 439
Db 378 accctgacatcagagacccttacaagcagctacgctggtgagagacccctgagatagc 437
QY 440 TCACCCAGCTAGCTCAACAAACCATGAGATCAGAGCTCGCAAACTCTTNTGACAAAG 499
Db 438 tcacccagctagctcaaaacccaatgagatcagagctcggcaaacctctctgtgacaag 497
QY 500 TCTTCGGGAGAGCGAGTCCCTGATGACGACATTTGTGATGCCATCAACCAAGTCTG 559
Db 498 tctctcggagaaagcggagctccctgaaatgacgacatgtgagctgcaacaacagcgtc 557
QY 560 ACTGCGGGGTATCCGCTGCTTCGTTATGAGATCAAGATATCCATGTGCCACCCCGG 619
Db 558 actgctgggtatccgctgctcctcgttataagatcagaatgatacctgtgtccaccccg 617
QY 620 TCAAAAGACTGTAGATGACAGTGGAGGACAGCGGAAAGGACGACAGTTCTAG 679
Db 618 tgaagagactatgacagatgacaggtgagagcagagcggcgaaacgggacacagttctag 677

QY 680 AGTCTGAGGGAGCCCGAGAGTCCGGCATATGATGGCAGAGGAAAGCAAGCCAGCA 739
Db 678 agtctgagggagcccgagagctcgccaatgtgagcagaaggaaacagccccaga 737
QY 740 TCTGTGCTTCGAAAGCAAGAAAGCTTCAACAGATTAATTCAGCAGCAGGAGCCAGTGC 799
Db 738 tctgtgcttcgcgaagcagaaagcttgaacagataaaatcagcagcagagagcagctg 797
QY 800 CAGTTTGGGAGAGGCCAAGCTTAAAGCTTAAGCTTATTCGAATCCTGGCTGCAGCTCTGA 859
Db 798 cagttctgggaaagggccaagctcaagctgaagctatcgaaatcccggtccgacctga 857
QY 860 CACAACATTAATGAGATGACAGACCTTCACTGACTGTGGCCGAGCAGATGTACGCCGT 919
Db 858 cacaacataatgagatgacagacttcaactgactgtggtgcgagcagtatgtcaagcgt 917
QY 920 TCTCCAACTGGCCAGAGACTCCAACTATCTACTGCCCCCTCAACCTTGGGATGTCA 979
Db 918 tctccaaactggccaaaggactccaacactatctactgcctccaacccctgcgaltga 977
QY 980 CCAGCATGTGGCTCAGGCCATGGGTATATGAGGCCCTCACCAAGCCCGAGTGCAG 1039
Db 978 ccagcatggtggtcagccagccatggtgtatataagagccctcaccaagccccagtgccag 1037
QY 1040 GGACTCCAGACTACTCTCCAGTGGAGCAGCAGAGATGTCACAGGATACAGATGCAAGT 1098
Db 1038 ggaactccagactcactctccagctggtgagcagcagagatgtccaggtgtacagatgcaagt 1096

RESULT 9

US-09-652-913-8809
; Sequence 8809, Application US/09652913
; GENERAL INFORMATION:
; APPLICANT: Fald, Dean R.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1174-001
; CURRENT APPLICATION NUMBER: US/09/652,913
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,107
; NUMBER OF SEQ ID NOS: 10833
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8809
; LENGTH: 1384
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-913-8809

Query Match 72.1%; Score 857; DB 25; Length 1384;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 140 GCCGCGCTCTCTGTGATTGCCCCGAAACACCGTGTACTGTTCTGTGCCGACAGAGAG 199
Db 138 gccgcgcctcctctgtatgctcccgaaacacgctgtactgtctgcgcgacagagag 197
QY 200 CCGGGTGTGGAGCCAAATGGCCGATTCACCGGATCCTGGAGCCCTGGTTGAACATCC 259
Db 198 cctcggtgtggtgagcgaatggtgacgtatccacggatccctggagccctgttgaacatcc 257
QY 260 TCATCCCTGTGTAGACCGGATCGATATGTGAGAGTCTCAAGAAATGTGATCAACG 319
Db 258 tcctccctgtgttagacggatccgatatgtgcagagctcccaagaaatgtgatacaag 317
QY 320 TGCCTGACAGTGGCTGTGACTCTCGACAAATGAATGATGAGTCCCTTT 379
Db 318 tgcctgagcagctgctgtgactctcgacaatgttaactctgcaaatcgatgagctctt 377
QY 380 ACCTGGCATGAGGACCCCTTACAAGCAAGCTACGCTGTGGAGGACCCCTGATATGCCG 439
Db 378 accctgacatcagagacccttacaagcagctacgctgtgagagacccctgagatagc 437

Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 140 GCCGGCCCTCTGATGATGCCCCGAAACAGCGTGCTACTGTTCTGCGCCGACAGAGG 199
    |||||||
Db 138 gccggccctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 197
QY 200 CCTGGGTGTGAGGCAATGAGGCCATTCCACCGGATCTCGAGCCTGTGTAACATCC 259
    |||||||
Db 198 cctgggtgtgaggaatgagggccattccacggatccttgagcctgtgttgaacatcc 257
QY 260 TCATCCCTGTGTAGACCGGATCCGATATGTCAGAGTCTCAGAGAAATTGTCATCAGC 319
    |||||||
Db 258 tcattccctgtgttagaccggtatcgtcagatgtcagagtcctcaagaaattgtcatcaacg 317
QY 320 TCCGTGAGAGTGGCTGTGACTCTCGACATGTAATCTGCAATCGATGAGTCCCTTT 379
    |||||||
Db 318 tgcctgagagcagtcgctgctgctgctgctgctgctgctgctgctgctgctgctgctt 377
QY 380 ACCTCGCGCATCATGAGACCTTACAGAGCAAGCTACGCTGTGAGAGACCTGTGATGCCG 439
    |||||||
Db 378 accctgcatactatgacaccttaacaaagctacggtgtgagagacctgtgtatgcccg 437
QY 440 TCACCCAGCTACTCAACACACCATGATGATCAGAGCTCGGCAACTCTTNTGACAAAG 499
    |||||||
Db 438 tcaccagactagctcaaaacacatgagatcagagctcggcaaaactctctgagcaaaag 497
QY 500 TCTTCCGGGAAGGGGAGTCCCTGTAATGCCAGATGTGATGCAATCAACAGAGCTGCTG 559
    |||||||
Db 498 tcttcgggaagaggagatccctggaatcgtgagatctggtgagtcacaaacagctgcgcg 557
QY 560 ACTGCTGGGGTATCCGCTCCCTNCCTTATGAGATCAAGAGATATCCATGTGCCACCCCGGG 619
    |||||||
Db 558 actgctgggtatccgctcgtcctcgltatgagatcaagatatccatgtgcccaccccggg 617
QY 620 TCAAGAGCTCTATGCAAGATGAGAGTGTGAGAGCAGAGCGGGGAAACGGCCATAGTTCTAG 679
    |||||||
Db 618 tgaagagactctatgagatgacagctgagagcagagcggagaaacgggccaacagctctag 677
QY 680 ACTCTGAGGGGAGCCGAGAGTGGCCATCAATGTGGCAGAGGAGAAAGAGGCCAGAGA 739
    |||||||
Db 678 agctcgaaggagaccgagagctgcatcaatgtgtgcagaagaggaaacgagcccaaga 737
QY 740 TCCCTGGCCTCCGAGCAGAGAAAGGCTGAACAGATAATAGCGCAGCAGAGAGGCCAGTGTG 799
    |||||||
Db 738 tccctggcctccgaagcagaagagctgaaacagataatcagagcagaagagggccaggtg 797
QY 800 CAGTTCTGGCGAAGCGCAAGGCTAAAGCTGAAGCTATTGGAATCCTGGCTGCAAGCTCTGA 859
    |||||||
Db 798 cagttctggcgaaagcgcaagctaaagctgaagctatctgaaatcctggtcgaagctctga 857
QY 860 CACAACATATGAGATGAGAGCAGAGCTTCACTGACTGTGGCCAGAGAGTATGTCAGCGCGT 919
    |||||||
Db 858 cacaacataatgagatgagatgagcagctctcaactgagctgtgagcagagctatgtcagagcgt 917
QY 920 TCTCCAACTGGCCCAAGAGCTCCAAACATATCTACTGCTCTCCAAACCTGTGCGATGTCA 979
    |||||||
Db 918 tctccaaactggcccaagagctccaaacatactctactgcccctccaaacccctgtgagatgca 977
QY 980 CCAGCATGTGTGCTCAGGCGCATGGGTATATGAGAGCCTTCACCAAGGCCCAAGTCCGAG 1039
    |||||||
Db 978 ccagcatgtgtgctcagcgcatgtgtatgtgagccctcccaagagcccaagctgtccag 1037
QY 1040 GGACATCAGACTCTCTCCAGTGGAGCAGCAGAGATGTCAGGGGTACAGATGCAAGT 1098
    |||||||
Db 1038 ggaactcagactcactctccaggtggagcagcagagatgttccagggtaacagatgtcaagt 1096
```

RESULT 7
US-09-652-127-7358
: Sequence 7358, Application US/09652127
: GENERAL INFORMATION:
: APPLICANT: Shyjan, Andrew W.

```
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES  
; FILE OF INVENTION: THEREFOR  
; FILE REFERENCE: 1600.1183-001  
; CURRENT APPLICATION NUMBER: US/09/652,127  
; PRIOR FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: 60/151,134  
; NUMBER OF SEQ ID NOS: 10475  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7358  
; LENGTH: 1384  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-652-127-7358
```

Query Match 72.1%; Score 857; DB 25; Length 1384;

Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 140 GCCGGCCCTCTGATGATGCCCCGAAACAGCGTGCTACTGTTCTGCGCCGACAGAGAG 199
    |||||||
Db 138 gccggccctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 197
QY 200 CCTGGGTGTGAGGCAATGAGGCCATTCCACCGGATCTCGAGCCTGTGTAACATCC 259
    |||||||
Db 198 cctgggtgtgaggaatgagggccattccacggatccttgagcctgtgttgaacatcc 257
QY 260 TCATCCCTGTGTAGACCGGATCCGATATGTCAGAGTCTCAGAGAAATTGTCATCAGC 319
    |||||||
Db 258 tcattccctgtgttagaccggtatcgtcagatgtcagagtcctcaagaaattgtcatcaacg 317
QY 320 TCCGTGAGAGTGGCTGTGACTCTCGACATGTAATCTGCAATCGATGAGTCCCTTT 379
    |||||||
Db 318 tgcctgagagcagtcgctgctgctgctgctgctgctgctgctgctgctgctgctgctt 377
QY 380 ACCTCGCGCATCATGAGACCTTACAGAGCAAGCTACGCTGTGAGAGACCTGTGATGCCG 439
    |||||||
Db 378 accctgcatactatgacaccttaacaaagctacggtgtgagagacctgtgtatgcccg 437
QY 440 TCACCCAGCTACTCAACACACCATGATGATGATCAGAGCTCGGCAACTCTTNTGACAAAG 499
    |||||||
Db 438 tcaccagactagctcaaaacacatgagatcagagctcggcaaaactctctgagcaaaag 497
QY 500 TCTTCCGGGAAGGGGAGTCCCTGTAATGCCAGATGTGATGCAATCAACAGAGCTGCTG 559
    |||||||
Db 498 tcttcgggaagaggagatccctggaatcgtgagatctggtgagtcacaaacagctgcgcg 557
QY 560 ACTGCTGGGGTATCCGCTCCCTNCCTTATGAGATCAAGAGATATCCATGTGCCACCCCGGG 619
    |||||||
Db 558 actgctgggtatccgctcgtcctcgltatgagatcaagatatccatgtgcccaccccggg 617
QY 620 TGAAGAGCTCTATGCAAGATGAGAGTGTGAGAGCAGAGCGGGGAAACGGCCACAGTCTGTG 679
    |||||||
Db 618 tgaagagctctatgagatgacagctgagagcagagcggagaaacgggccaacagctctag 677
QY 680 AGTCTGAGGGGAGCCGAGAGTGGCCATCAATGTGGCAGAGGAGAAAGAGGCCAGAGA 739
    |||||||
Db 678 agtctgagggagaccgagagctgcatcaatgtgtgcagaagaggaaacagggcccaaga 737
QY 740 TCCCTGGCCTCCGAGCAGAGAAAGGCTGAACAGATAATAGCGCAGCAGAGAGGCCAGTGTG 799
    |||||||
Db 738 tccctggcctccgaagcagaagagctgaaacagataatcagagcagaagagggccaggtg 797
QY 800 CAGTTCTGGCGAAGCGCAAGGCTAAAGCTGAAGCTATTGGAATCCTGGCTGCAAGCTCTGA 859
    |||||||
Db 798 cagttctggcgaaagcgcaagctaaagctgaagctatctgaaatcctggtcgaagctctga 857
QY 860 CACAACATATGAGATGAGAGCAGAGCTTCACTGACTGTGGCCAGAGAGTATGTCAGCGCGT 919
    |||||||
Db 858 cacaacataatgagatgagatgagcagctctcaactgagctgtgagcagagctatgtcagagcgt 917
QY 920 TCTCCAACTGGCCCAAGAGCTCCAAACATATCTACTGCTCTCCAAACCTGTGCGATGTCA 979
    |||||||
```

```

Db 618 tgaagagctatgcaagatgcaagtggaagcagcgcggaacagcgccacagtlctag 677
OY 680 AGACTGAGGGAGCCCGAGTCCGACCTCATGTCGACGAGAGGAACAAAGCCCGAGA 739
Db 678 agctcgaaggaacccgaagatcgccatcaatgugcagaagaagaaacaaagcccgaga 737
OY 740 TCTGCGCTCCGAGCAGAAAGGCTGACAGATTAATTCAGGACGACGAGAGCCAGTG 799
Db 738 tcttgccctcgaagcagaagaagctgaacagblaaatcgaagcagcgagagcgccagtg 797
OY 800 CAATTCTGGGGAAGGCCAAGCTTAAGCTAAGCTATTGCAATTCCTGCTGCAGCTTGA 859
Db 798 cagltcggcgaaagccaaagctaaagctgaatctcgaatccctggcgcagctctga 857
OY 860 CACAACATATGAGATGACGACGCTTCAGTACTGCTGCGCGAGCAATATGTACGCGGT 919
Db 858 caacaataatgagatgacgacagctcactgacttgcgcgaagcaatgtaagcggt 917
OY 920 TCTCCAACTGGCCCAAGACTCCACACTATCTACTGCCCTCCAAACCTGGCGATGTC 979
Db 918 tctccaactgcccgaagactccacaactatctcactgcccctccaaacctgycagatgca 977
OY 980 CCAGCATGTGCTGCTCAGGCGCATGGGTATATGAGAGCCCTCACCAGAGCCCGAGTCC 1039
Db 978 ccagcatggtgctcagcgccatggtatataagagccctcacaagaagccccagtgccag 1037
OY 1040 GGACTCCAGACTCACTCTCCAGTGGAGAGCAGAGATGTCAGAGGTACAGATGCAAGT 1098
Db 1038 ggaactccagactcactctccagtggaagcagcagagatgctccagggtaacagatgcaagt 1096

```

RESULT 5

```

US-09-649-162-7309
; Sequence 7309, Application US/09649162
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; APPLICANT: Richardson, Jennifer
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1181-001
; CURRENT APPLICATION NUMBER: US/09/649,162
; CURRENT FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: 60/151,057
; PRIOR FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 9990
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 7309
; LENGTH: 1384
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-649-162-7309

```

Query Match 72.1%; Score 857; DB 25; Length 1384;

```

Best Local Similarity 99.8%; Pred. No. 0;
Matches 957; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 140 GCCGCGCTCTCTGATTGCCCGAAGACGCTGACTGTTGCTGCCGACGAGAGG 199
Db 138 gccgcgctctctctgattgccccgaacacgcgtgtaactgcttcgacagcaggaag 197
OY 200 CCGGGGAGGAGGAGGAATGGGCGGCTTCACGCGATCCCTGGAGCCCTGGTTAACTCC 259
Db 198 cctgggctgggaggaatgggcccgaattccacgcgaltccggagccctgggttaaatcc 257
OY 260 TCATCCCTGTGTAGACCGGATCGATATGTGAGAGTCTCAAGAAATGTATCAAG 319
Db 258 tcatccctgtgttagaccggatccgatatgtcagaggtctcagaagaaatgtcatcaacg 317
OY 320 TGCCTGAGCAGTGGCTGTGACTCTCGACAATTAACCTGCAAAATCGATGAGTCTTT 379

```

```

Db 318 tgcctgagcagtcggtctgactctcgacaatglaactctgcaaatcgatgagctctt 377
OY 380 ACCTGCCATCATGAGACCCCTTACAGCAAGCTACGGTGTGGAGGACCCGAGATGCGG 439
Db 378 acctgccaatcagacccttacaagcaagctacggttggaggaacccgaatgacg 437
OY 440 TCAOCCAGCTAGCTCAACCAACCATGAGATCAGAGCTCGGCAACCTCTTNTGACAAAG 499
Db 438 tcaocccagctagctcaacaacatgagatcagagctcgcgcgaactctctcagacaag 497
OY 500 TCTTCCGGGAAGGAGTCCCTGAATGCCAGCATTTGATGCCATCAACCAAGTCTG 559
Db 498 tcttccgggaacggaagctccctgaatgcccagcatltggaatgcataaaccaagctgctg 557
OY 560 ACTGCTGGGTATCCGCTGCTCCTGCTGCTATGATGATCAAGGATATTCATGTCGCCCGGG 619
Db 558 actgctgggtatccgctctgctcctcgttataagatcaagatcccatggtccacccggg 617
OY 620 TGAAGAGTCTATGACAGATGAGTGAGGACAGAGCGCGGAAACGGCCACAGTTCTAG 679
Db 618 tgaagagctatgcaagatgcaagtggaagcagagcgcggaacggyccacagtlctag 677
OY 680 AGTCTGAGGGAGCCCGAGAGTCCGCCATCAATGTGCGAAGAGGAACACAGCCGAGA 739
Db 678 agtctgagggagcccgagagctcgccatcaatgltggcagaagaagaaacagggccaga 737
OY 740 TCTGCGCTCCGAGCAGAAAGGCTGAAACAGATTAATTCAGGACGACGAGAGCCAGTG 799
Db 738 tcttgccctcgaagcagaagaagctcgaacagatbaatccagcagaagagggccagtg 797
OY 800 CAATTCTGGGGAAGGCCAAGGCTAAAGCTTAAGCTATTGCAATCTTGCTGCAGCTCTGA 859
Db 798 cagltcggcgaaagccaaagctaaagctgaagctatctgaatccctgycgcagctctga 857
OY 860 CACAACATATGAGATGACGACGCTTCACTGACTGTGGCGGACGAGTATGTACGCGGT 919
Db 858 caacaataatgagatgacgacagctcactgacttgcgcgaagctatgtaagcggt 917
OY 920 TCTCCAACTGGCCCAAGACTCCACACTATCTACTGCCCTCCAAACCTGGCGAGTGA 979
Db 918 tctccaactgcccgaagactccacaactatctcactgcccctcacaacctggcgatgca 977
OY 980 CCAGCATGTGCTGCTCAGGCGCATGGGTATATGAGAGCCCTCACCAGAGCCCGAGTCC 1039
Db 978 ccagcatggtgctcagcgccatggtatataagagccctcacaagaagccccagtgccag 1037
OY 1040 GGACTCCAGACTCACTCTCCAGTGGAGAGCAGAGATGTCAGAGGTACAGATGCAAGT 1098
Db 1038 ggaactccagactcactctccagtggaagcagcagagatgctccagggtaacagatgcaagt 1096

```

RESULT 6

```

US-09-652-109-7925
; Sequence 7925, Application US/09652109
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1180-001
; CURRENT APPLICATION NUMBER: US/09/652,109
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,128
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 10105
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 7925
; LENGTH: 1384
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-109-7925

```

Query Match 72.1%; Score 857; DB 25; Length 1384;


```

: TITLE OF INVENTION: HUMAN PROSTATE STROMAL LIBRARY
:
: FILE REFERENCE: MLN98-25P
:
: CURRENT APPLICATION NUMBER: US/09/338,467
:
: CURRENT FILING DATE: 1999-06-22
:
: EARLIER APPLICATION NUMBER: 60/090,177
:
: EARLIER FILING DATE: 1998-06-22
:
: NUMBER OF SEQ ID NOS: 1022
:
: SOFTWARE: FASTSEQ for Windows Version 3.0
:
: SEQ ID NO 941
:
: LENGTH: 1384
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: OS-09-338-467-941

```

Query Match	72.1%	Score 857;	DB 17;	Length 1384;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 957;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

Oy	140	GCCTGGCTCTCTCGGATGTTGGCCCGAACAACCGGCGATCGTTCTGCGCCACACAGAG	199
Db	138	gcgcgcctctctcttgatctgcccgcgaacacgctgtaactcttcgctgcgcgaagag	197
Oy	200	CCTGGGTGTGTGACCGAATGGCGCGATTCCACCGATCTCGAGACCGCTGGTTTGAATCC	259
Db	198	ccctggctgctgagcgaatgggacgttccaccggatccctggagcctggtttbaacatcc	257
Oy	260	TCATCCCTGTGTAGACCGGATCCGATATGTGCGAGTCTCAAGGAATTTGTTCATACG	319
Db	258	tcatccctctgttagaccggatccgatctatgycagagctccaaagaaattgtcatcaag	317
Oy	320	TGCTTGACACAGTCCGCTGTGACTTTCGACATGTATACCTGTGCAATCGATGAGTCTTT	379
Db	318	tgccctgagcagctgcgctgtgactcttcgacaatgttaactctgcaatcgatgtagtccctt	377
Oy	380	ACCTCCGATCATGGACCCCTTACAAGCAAGCTACGGTGTGGAGGACCTCTAGATGCGC	439
Db	378	acctgcgacatcaggaaccttaacaggcaagctacggtctggaagacctgtagatgcg	437
Oy	440	TCACCCACATGACCTCAAAACAACATGAGATAGAGCTCGCGCAACCTCTCNTGGACAAG	499
Db	438	tcaaccagctatgcttcaaacacacatctgatacgctccggaacatctctcttgacaaag	497
Oy	500	TCCTTCGGGGAACGGGAGTCCCTGATGCGACAGATTGTGATGCCATCAACCAAGTGTCTG	559
Db	498	ctctccggaacgcggagctccctgtaatgccagcatctgtgatatgcatcaacaaagctgcg	557
Oy	560	ACTGCTGGGGGATATCCGCTGCTNGCTTATGAGATCAAGGATATCCATGTGCCACCCGGG	619
Db	558	actgtctgggtatctccgctcgcctcgttatgataccaagatataccaatgctccaccgcgg	617
Oy	620	TGAAAGACTCTATGACAGATGCGAGTGGAGTGGAGCGCGGCAAAACGGGGCCACAGTTTAG	679
Db	618	tgaagaagctcatgccaagctcgaagctggaaggaagagcgcggaataacggtcttcag	677
Oy	680	AGTCTGAGGGGACCCGAGAGTCCGGCCATCATGTGTGGCAGAAAGGGAAGAAACAGGCCAGA	739
Db	678	agctcgaggggaacccgagagctcgccaatactgtgcaaggaaggaagaataacagcccaaga	737
Oy	740	TCCTTGCCCTCCGAGACGAAAGGCTGAACAGATPAATCAGGCGACGAGAGGCCAGTG	799
Db	738	tccttgccctccgaagcagaataaagctgacagataaatacaagcagaggaagccagtg	797
Oy	800	CAGTCTTGCGGCAAGGCCAAGGCTTAAAGCTGAAGCTATTCGAATCCTGTGCGACGCTTGA	859
Db	798	cagttctctgcgaaggcgaagctlaaagcttgaagctatctcgaaatccctgcgcagctctga	857
Oy	860	CACAACATPAATGGAATGCACAGCTTACAGTACGAGCTGTGGCCGACAGATATCTAGCCGT	919
Db	858	cacaacataatgagatgctcagcagcttcaactcgtactgtggtccgagcagtatgtcagcgct	917
Oy	920	TCTCCAACTGGCCAAGAGCTCCAACTATCTACTGCTCCCAACCTTGCGGATGTCA	979

Db 918 tctccaacatgccaagactccaacatactctactagtcctccaacactgysatg tca 977

QY 980 CCAGCATGTTGGCTTAGGCCCATGGGTATATATGCAAGCCTTACCAAAAGCCCACTGGCAG 1038

Db 978 ccagatctgtgctcaaggccattggatglatatgagccctaccaaaagcccacgtgcag 1038

QY 1040 GGACTCCAGACTCACTCTCCAGTGGGAGACAGAGATGTCACAGGATACAGATGCAAGT 1098

Db 1038 ggaactccagactcaactctccagctggagacagagatagtcctcagggatcacatgcaat 1096

RESULT 4
US-09-644-869-8024

```

1 GENERAL INFORMATION:
2 APPLICANT: Shyjan, Andrew W.
3 APPLICANT: McCarthy, Sean A.
4 APPLICANT: Holtzman, Douglas A.
5 APPLICANT: Monahan, John
6 APPLICANT: Richardson, Jennifer
7 TITLE OF INVENTION: NOVEL NUCLEIC ACID MO
8 TITLE OF INVENTION: THEREFOR
9 FILE REFERENCE: 1600.1182-001
10 CURRENT APPLICATION NUMBER: US/09/644,869
11 CURRENT FILING DATE: 2000-08-28
12 PRIOR APPLICATION NUMBER: 60/151,062
13 PRIOR FILING DATE: 1999-08-27
14 NUMBER OF SEQ. ID NOS: 9708
15 SOFTWARE: FASTSEQ for Windows Version 4.0.
16 SEQ. ID NO 8024
17 LENGTH: 1384
18 TYPE: DNA
19 ORGANISM: Homo sapiens
20 US-09-644-869-8024

```

Query Match	72.1%;	Score 857;	DB 25;	Length 1384;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 957;	Conservative	0;	Mismatches 2;	Indels 0;
				Gaps 0;
QY 140	GGCGGCGCTCCTGTGATTTGCCCGGAACACCGTGTACTGTTCTGCGGACGACGAGG	199		
DB 138	ggcgcgccctccttgatctgcccgaacaacgltgtaactgltctgcccgaacgaagaag	197		
QY 200	CCGTGGTGTGTGAGCAAAATGGGCGCAATTCACAGATTCCTGTGAGCCCTGGTTTGAAATCC	259		
DB 198	cctgggtgtgtgagcgaaatggygcggtctcacccgagctctgagcctgtttgaacatcc	257		
QY 260	TCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGAAATTTGTATCAACG	319		
DB 258	tcatccctgtgttagacggtatctgatatgtgagaagcttcaagaatatgtatcatacg	317		
QY 320	TGCGTTAGACAGTGCCTGTGTGACTCTTGACAAATCTACTGCAAAATCGATGGAGTCTTT	379		
DB 318	tgccttagagagtgtgctgtgtactctgcgaacaatgtaactctgcgaatctgagtgcctt	377		
QY 380	ACCTGGCATCATGAGACCTTTACAAAGCAAGCTACGCTGTGAGAGACCTCGAATATCCG	439		
DB 378	acctggtcatcatgagaccccttaacaagcaagctacggttgtgagagccctgtgatacg	437		
QY 440	TCACCCAGTAGCTCAACAAACCATGAGATCAGAGCTCGGCAAACTCTCTNTGGACAAG	499		
DB 438	tcaaccagctagcttcaacaacatgagatcagatctcgcaaacctctcttgacaag	497		
QY 500	TCCTTCGGGGAACGGAGTCCCTGATATGCCAGCATTTGTGATGTGCATCAACCAAGTCTTG	559		
DB 498	tcttcctgggaacggaggtcccttgaatgcagcatgtgtgatagcatacaacaagctgcgtg	557		
QY 560	ACGTGTGGGGTATCCCTGCTGCTNCGTTTAGATGATCAAGGATATCCATATGCCACCCGGG	619		
DB 558	acgtgtgggtataccgtctgcctcccttataagatcaagaatatcaatctgtgccaccccggtg	617		
QY 620	TGAAGAAGTCTATGCAATGTCAGGTGGAGGCAAGCGCGGGAACGGCGCACATTTCTAG	679		

```

1  PRIOR FILING DATE: 1997-06-13
2  PRIOR APPLICATION NUMBER: 60/049,550
3  PRIOR FILING DATE: 1997-06-13
4  PRIOR APPLICATION NUMBER: 60/049,566
5  PRIOR FILING DATE: 1997-06-13
6  PRIOR APPLICATION NUMBER: 60/049,606
7  PRIOR FILING DATE: 1997-06-13
8  PRIOR APPLICATION NUMBER: 60/049,607
9  PRIOR FILING DATE: 1997-06-13
10 PRIOR APPLICATION NUMBER: 60/049,608
11 PRIOR FILING DATE: 1997-06-13
12 PRIOR APPLICATION NUMBER: 60/049,609
13 PRIOR FILING DATE: 1997-06-13
14 PRIOR APPLICATION NUMBER: 60/049,610
15 PRIOR FILING DATE: 1997-06-13
16 PRIOR APPLICATION NUMBER: 60/049,611
17 PRIOR FILING DATE: 1997-06-13
18 PRIOR APPLICATION NUMBER: 60/050,901
19 PRIOR FILING DATE: 1997-06-13
20 PRIOR APPLICATION NUMBER: 60/052,989
21 PRIOR FILING DATE: 1997-06-13
22 PRIOR APPLICATION NUMBER: 60/051,919
23 PRIOR FILING DATE: 1997-07-08
24 PRIOR APPLICATION NUMBER: 60/055,984
25 PRIOR FILING DATE: 1997-08-18
26 PRIOR APPLICATION NUMBER: 60/058,665
27 PRIOR FILING DATE: 1997-09-12
28 PRIOR APPLICATION NUMBER: 60/058,668
29 PRIOR FILING DATE: 1997-09-12
30 PRIOR APPLICATION NUMBER: 60/058,669
31 PRIOR FILING DATE: 1997-09-12
32 PRIOR APPLICATION NUMBER: 60/058,750
33 PRIOR FILING DATE: 1997-09-12
34 PRIOR APPLICATION NUMBER: 60/058,971
35 PRIOR FILING DATE: 1997-09-12
36 PRIOR APPLICATION NUMBER: 60/058,972
37 PRIOR FILING DATE: 1997-09-12
38 PRIOR APPLICATION NUMBER: 60/058,975
39 PRIOR FILING DATE: 1997-09-12
40 PRIOR APPLICATION NUMBER: 60/060,834
41 PRIOR FILING DATE: 1997-10-02
42 PRIOR APPLICATION NUMBER: 60/060,841
43 PRIOR FILING DATE: 1997-10-02
44 PRIOR APPLICATION NUMBER: 60/060,844
45 PRIOR FILING DATE: 1997-10-02
46 PRIOR APPLICATION NUMBER: 60/060,865
47 PRIOR FILING DATE: 1997-10-02
48 PRIOR APPLICATION NUMBER: 60/061,059
49 PRIOR FILING DATE: 1997-10-02
50 PRIOR APPLICATION NUMBER: 60/061,060
51 PRIOR FILING DATE: 1997-10-02
52 NUMBER OF SEQ ID NOS: 737
53 SOFTWARE: Patentin Ver. 2.0
54 SEQ ID NO 45
55 LENGTH: 1337
56 TYPE: DNA
57 ORGANISM: Homo sapiens
58 FEATURE:
59 NAME/KEY: SITE
60 LOCATION: (1335)
61 OTHER INFORMATION: n equals a,t,g, or c
62 NAME/KEY: SITE
63 LOCATION: (1336)
64 OTHER INFORMATION: n equals a,t,g, or c
65 NAME/KEY: SITE
66 LOCATION: (1337)
67 OTHER INFORMATION: n equals a,t,g, or c
68 IS-09-209-462B-45

```

Query Match	72.1%;	Score 857;	DB 16;	Length 1337;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 957;	Conservative	0;	Mismatches 2;	Indels 0;
				Gaps 0;

[illegible]

```

RESULT      3
US-09-338-467-941
: Sequence 941, Application US/09338467
: GENERAL INFORMATION:
: APPLICANT: Gearing, David P.
: APPLICANT: Holtzman, Douglas A.
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A

```

c	37	572	48.1	1309	17	US-09-359-922-11863	Sequence 11863, A
c	38	572	48.1	1309	17	US-09-359-922-11863	Sequence 11863, A
c	39	509	42.8	622	36	US-09-998-998-1161	Sequence 1161, A
c	40	454	36.2	454	33	US-09-878-178-193	Sequence 193, App
c	41	454	38.2	454	37	US-10-046-935-193	Sequence 193, App
c	42	453	38.1	546	37	US-10-081-124-244	Sequence 244, App
c	43	453	38.1	550	1	PCT-US02-02870-2995	Sequence 2995, Ap
c	44	453	38.1	550	37	US-10-066-543-2995	Sequence 2995, Ap
c	45	417	35.1	483	25	US-09-652-127-341	Sequence 341, App
c	32	793	66.8	1416	22	US-09-598-075A-249	Sequence 249, App
c	33	728	67.3	1429	55	US-60-172-773-10483	Sequence 10483, A
c	34	674	56.7	1234	55	US-60-164-285-3468	Sequence 3468, Ap
c	35	582	49.0	732	33	US-09-878-134-349	Sequence 349, App
c	36	582	49.0	732	37	US-09-033-356-349	Sequence 349, App
c	37	572	48.1	1309	17	US-09-359-922-11863	Sequence 11863, A
c	38	572	48.1	1309	17	US-09-359-922-11863	Sequence 11863, A
c	39	509	42.8	622	36	US-09-998-998-1161	Sequence 1161, A
c	40	454	36.2	454	33	US-09-878-178-193	Sequence 193, App
c	41	454	38.2	454	37	US-10-046-935-193	Sequence 193, App
c	42	453	38.1	546	37	US-10-081-124-244	Sequence 244, App
c	43	453	38.1	550	1	PCT-US02-02870-2995	Sequence 2995, Ap
c	44	453	38.1	550	37	US-10-066-543-2995	Sequence 2995, Ap
c	45	417	35.1	483	25	US-09-652-127-341	Sequence 341, App

ALIGNMENTS

```

PCT-US98-12125-45
; Sequence 45, Application PC/TUS9812125
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 86 Human Secreted Proteins
NUMBER OF SEQUENCES: 318
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US98/12125
FILING DATE: Jan 01, 1990
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PZ008Complete
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 1337 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
PCT-US98-12125-45

```

	Query Match	Best Local Similarity	Score	DB 1	Length
Matches	957	Conservative	99.8%	Pred. No. 0	Mismatches 2; Indels 0; Gaps 0
QY	140	GGCGGGCCCTCCCTTGATTTGCCCGAAGCAGCTGTACTGTTGCGCGGACAGAGAG	199		
Db	152	GGCGGGCCCTCCTCTGATGATGCCCCGGAACACCTGTACTGTTGCGCGGACAGAGAG	211		
QY	200	CCTGGGTGTGAGAGCGAATGGGCCGATTCACCGGATCCTGGAGCCTGTTTGAACATCC	259		
Db	212	CCTGGGTGTGAGAGCGAATGGGCCGATTCACCGGATCCTGGAGCCTGTTTGAACATCC	271		

QY	260	TCACCCCTGCTTAGACCGGATCCGATATGTGCACAGTCTCAAGGAAATTTCTCATCAAG	319
Db	272	TCATCCCTGTGTTAGACCGGATCCGATATGTGCACAGTCTCAAGGAAATTTCTCATCAAG	331
QY	320	TGACCTGAGCAGTGGGCTGTGACTCTTCGACATATGTAACCTCTGCAAAATCGATGGAGTCCTTT	379
Db	332	TGCGGTGAGCAGTGGGCTGTGACTCTTCGACATATGTAACCTCTGCAAAATCGATGGAGTCCTTT	391
QY	380	ACCTGGGCATCATGAGACCTTTACAAAGGCAAGCTACGGGTGGAGGAGCCCTGAGTATGGC	439
Db	392	ACCTGGGCATCATGAGACCTTTACAAAGGCAAGCTACGGGTGGAGGAGCCCTGAGTATGGC	451
QY	440	TCACCCAGCTAGCTCAAAACAACCATGAGATTCAGAGCTCGGCAAACTCTCTTNGACAAG	499
Db	452	TCACCCAGCTAGCTCAAAACAACCATGAGATTCAGAGCTCGGCAAACTCTCTCGGACAAG	511
QY	500	TCCTTCGGGAGACGGGAGTCCCTGATGCCAGCATTTGTGATGCCATCAACAAGCTGCTG	559
Db	512	TCCTTCGGGAGACGGGAGTCCCTGATGCCAGCATTTGTGATGCCATCAACAAGCTGCTG	571
QY	560	ACCTGCTGGGGATTCGGGTGCTTCGTTATGAGTCAAGGATATTCATGTGCCACCCCGGG	619
Db	572	ACCTGCTGGGGATTCGGGTGCTTCGTTATGAGTCAAGGATATTCATGTGCCACCCCGGG	631
QY	620	TGAAGAGTATTCAGAGATCAGGTGGAGGACAGCGCGGGAAGCGGGCACATTTCTAG	679
Db	632	TGAAGAGTATTCAGAGATCAGGTGGAGGACAGCGCGGGAAGCGGGCACATTTCTAG	691
QY	680	AGTCTGAGGGGACCCGAGAGTCCGGCATCATATGTGCACAGAAGGAAAGAACAGGCCACGA	739
Db	692	AGTCTGAGGGGACCCGAGAGTCCGGCATCATATGTGCACAGAAGGAAAGAACAGGCCACGA	751
QY	740	TCCTGGGCTCCGAAGCGAAGAAAGGCTGAACAGATTAATCAGGACGAGAGAGGCCAGTG	799
Db	752	TCCTGGGCTCCGAAGCGAAGAAAGGCTGAACAGATTAATCAGGACGAGAGAGGCCAGTG	811
QY	800	CAGTTCTGGCGAAGGCCAAGGCTTAAAGCTGAAGCTATTCGATCTGCTGCAGCTCTGA	859
Db	812	CAGTTCTGGCGAAGGCCAAGGCTTAAAGCTGAAGCTATTCGATCTGCTGCAGCTCTGA	871
QY	860	CACAACATAATGAGATGAGACAGCTTCACTGACTGGCGCGAGAGTATGTACAGCGCT	919
Db	872	CACAACATAATGAGATGAGACAGCTTCACTGACTGGCGCGAGAGTATGTACAGCGCT	931
QY	920	TCTCCAAACTGGCCAGAGGACTCCAAACACTATCTTACTGCTTCCAAACCTTGGCGATGTCA	979
Db	932	TCTCCAAACTGGCCAGAGGACTCCAAACACTATCTTACTGCTTCCAAACCTTGGCGATGTCA	991
QY	980	CCAGCATGTGGCTCAGGCGCATGGGTTATATGAGAGCCCTCACCAAAGCCCAAGTCCAG	1033
Db	992	CCAGCATGTGGCTCAGGCGCATGGGTTATATGAGAGCCCTCACCAAAGCCCAAGTCCAG	1051
QY	1040	GGACATCCAGACTCACTCTCCAGTGGGAGCAGAGAGATGTCAGGGTTAAGATGCAAGT	1098
Db	1052	GGACATCCAGACTCACTCTCCAGTGGGAGCAGAGAGATGTCAGGGTTAAGATGCAAGT	1110

2
US-09-209-462B-45
Sequence 45, Application US/09209462B
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 86 Human Secreted Protein
FILE REFERENCE: P2008P1
CURRENT APPLICATION NUMBER: US/09/209,462
CURRENT FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: PCT/US98/12125
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/049,547
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,548
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,549

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 16:20:29 : Search time 2869.55 Seconds
(without alignments)
8956.896 Million cell updates/sec

Title: US-09-898-216-2

Perfect score: 1188
Sequence: 1 GGCCTTCGGAGCNAACCGCT.....GGAGACAGATTTCCTGATT 1188

Scoring table: OLIGO_NUC
Gapco 60.0 , Gapext 60.0

Searched: 21979536 segs, 10817449327 residues

Word size : 0

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Pending_Patents_NA_Main: *
1: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq: *
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq: *
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq: *
4: /cgn2_6/ptodata/1/pna/US080_COMB.seq: *
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq: *
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq: *
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq: *
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq: *
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq: *
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq: *
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq: *
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq: *
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq: *
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq: *
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq: *
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq: *
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq: *
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq: *
19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq: *
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq: *
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq: *
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq: *
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq: *
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq: *
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq: *
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq: *
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq: *
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq: *
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq: *
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq: *
31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq: *
32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq: *
33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq: *
34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq: *
35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq: *
36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq: *
37: /cgn2_6/ptodata/1/pna/US100_COMB.seq: *
38: /cgn2_6/ptodata/1/pna/US101_COMB.seq: *
39: /cgn2_6/ptodata/1/pna/US600_COMB.seq: *
40: /cgn2_6/ptodata/1/pna/US6001_COMB.seq: *
41: /cgn2_6/ptodata/1/pna/US6002_COMB.seq: *
42: /cgn2_6/ptodata/1/pna/US6003_COMB.seq: *
43: /cgn2_6/ptodata/1/pna/US6004_COMB.seq: *

44: /cgn2_6/ptodata/1/pna/US6005_COMB.seq: *
45: /cgn2_6/ptodata/1/pna/US6006_COMB.seq: *
46: /cgn2_6/ptodata/1/pna/US6007_COMB.seq: *
47: /cgn2_6/ptodata/1/pna/US6008_COMB.seq: *
48: /cgn2_6/ptodata/1/pna/US6009_COMB.seq: *
49: /cgn2_6/ptodata/1/pna/US6010_COMB.seq: *
50: /cgn2_6/ptodata/1/pna/US6011_COMB.seq: *
51: /cgn2_6/ptodata/1/pna/US6012_COMB.seq: *
52: /cgn2_6/ptodata/1/pna/US6013_COMB.seq: *
53: /cgn2_6/ptodata/1/pna/US6014_COMB.seq: *
54: /cgn2_6/ptodata/1/pna/US6015_COMB.seq: *
55: /cgn2_6/ptodata/1/pna/US6016_COMB.seq: *
56: /cgn2_6/ptodata/1/pna/US6017_COMB.seq: *
57: /cgn2_6/ptodata/1/pna/US6018_COMB.seq: *
58: /cgn2_6/ptodata/1/pna/US6019_COMB.seq: *
59: /cgn2_6/ptodata/1/pna/US6020_COMB.seq: *
60: /cgn2_6/ptodata/1/pna/US6021_COMB.seq: *
61: /cgn2_6/ptodata/1/pna/US6022_COMB.seq: *
62: /cgn2_6/ptodata/1/pna/US6023_COMB.seq: *
63: /cgn2_6/ptodata/1/pna/US6024_COMB.seq: *
64: /cgn2_6/ptodata/1/pna/US6025_COMB.seq: *
65: /cgn2_6/ptodata/1/pna/US6026_COMB.seq: *
66: /cgn2_6/ptodata/1/pna/US6027_COMB.seq: *
67: /cgn2_6/ptodata/1/pna/US6028_COMB.seq: *
68: /cgn2_6/ptodata/1/pna/US6029_COMB.seq: *
69: /cgn2_6/ptodata/1/pna/US6030_COMB.seq: *
70: /cgn2_6/ptodata/1/pna/US6031_COMB.seq: *
71: /cgn2_6/ptodata/1/pna/US6032_COMB.seq: *
72: /cgn2_6/ptodata/1/pna/US6033_COMB.seq: *
73: /cgn2_6/ptodata/1/pna/US6034_COMB.seq: *
74: /cgn2_6/ptodata/1/pna/US6035_COMB.seq: *
75: /cgn2_6/ptodata/1/pna/US6036_COMB.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	857	72.1	1337	1 PCT-US98-12125-45	Sequence 45, Appl
2	857	72.1	1337	16 US-09-209-462B-45	Sequence 45, Appl
3	857	72.1	1384	17 US-09-338-467-941	Sequence 941, App
4	857	72.1	1384	25 US-09-644-869-8024	Sequence 8024, App
5	857	72.1	1384	25 US-09-640-163-7309	Sequence 7309, App
6	857	72.1	1384	25 US-09-652-109-7925	Sequence 7925, App
7	857	72.1	1384	25 US-09-652-127-7358	Sequence 7358, App
8	857	72.1	1384	25 US-09-652-816-7419	Sequence 8809, App
9	857	72.1	1384	25 US-09-652-913-8809	Sequence 8809, App
10	857	72.1	1384	25 US-09-652-914-7997	Sequence 7997, App
11	857	72.1	1384	31 US-09-808-384-941	Sequence 941, App
12	857	72.1	1398	1 PCT-US01-03800A-1001	Sequence 1001, App
13	857	72.1	1398	1 PCT-US01-04058A-2212	Sequence 2212, App
14	857	72.1	1398	18 US-09-471-275-865	Sequence 865, App
15	857	72.1	1398	18 US-09-496-914A-8410	Sequence 8410, App
16	857	72.1	1398	22 US-09-560-875A-8410	Sequence 8410, App
17	857	72.1	1501	16 US-09-298-733-35	Sequence 35, Appl
18	857	72.1	1501	16 US-09-298-733A-35	Sequence 35, Appl
19	857	72.1	1501	29 US-09-723-594-35	Sequence 35, Appl
20	857	72.1	1501	29 US-09-724-497-35	Sequence 35, Appl
21	857	72.1	9098	1 PCT-US01-04942A-34	Sequence 34, Appl
22	857	72.1	9098	22 US-09-596-193A-37	Sequence 37, Appl
23	857	72.1	9278	1 PCT-US01-08656-5097	Sequence 5097, App
24	855	72.0	1322	1 PCT-US98-12125-102	Sequence 102, App
25	855	72.0	1322	16 US-09-209-462B-106	Sequence 106, App
26	794	66.8	1308	16 US-09-205-070-13379	Sequence 13379, App
27	794	66.8	1308	17 US-09-340-623-13379	Sequence 13379, App
28	794	66.8	1308	33 US-09-898-888-13379	Sequence 13379, App
29	793	66.8	1308	33 US-09-898-888A-13379	Sequence 13379, App
30	793	66.8	1416	1 PCT-US01-04098A-244	Sequence 244, App
31	793	66.8	1416	22 US-09-598-075-249	Sequence 249, App

Best Local Similarity 100.08; Pred. No. 44;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 554 CTGCTGACTGCTGGGT 570
 |||||
 Db 553 CTGCTGACTGCTGGGT 569

Search completed: September 22, 2002, 17:24:48
 Job time: 7095 sec

```

Oy 1 GGCTTCTGGAGCANNACCGCTCCGCTGCTCTGTTGGTTCGGAGGTCGTGCGGGCGTGG 60
Db 1 GGGTCTTCTGGAGCANNACCGCTCCGCTGCTCTGTTGGTTCGGAGGTCGTGCGGGCGTGG 60
Oy 61 GAAATGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 61 GAAATGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 119
Oy 121 GGGCTTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
Db 120 GGGCTTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
Oy 180 GTTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 239
Db 180 GTTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 239
Oy 240 GGAGCCTGTTGGAACATCTCATCTCCCTGTGTAGACCGGATCCGATATGTGACAGATCT 299
Db 240 GGAGCCTGTTGGAACATCTCATCTCCCTGTGTAGACCGGATCCGATATGTGACAGATCT 299
Oy 300 CAAGGAATTTCTCATCAACGTCGCTGACACATCGGCTGTGACTCTCGCAAAATGTAATCT 359
Db 300 CAAGGAATTTCTCATCAACGTCGCTGACACATCGGCTGTGACTCTCGCAAAATGTAATCT 359
Oy 360 GCAATATCGATGGAGTCCTTCTACCTGCGCATGTGAGACCCCTTACAAAGCAAGCTACGGTGT 419
Db 360 GCAATATCGATGGAGTCCTTCTACCTGCGCATGTGAGACCCCTTACAAAGCAAGCTACGGTGT 419
Oy 420 GGAGGACCTGATATGCGCGTACCCAGCTAGCTGCTCAAAACCAACATGAGATCGAGCTCGG 479
Db 420 GGAGGACCTGATATGCGCGTACCCAGCTAGCTGCTCAAAACCAACATGAGATCGAGCTCGG 479
Oy 480 CAACTCTCTTGTGACAAAGCTTTCGGGAAAGGAGTCCCTGAAATGCCAGATGTGGA 539
Db 480 CAACTCTCTTGTGACAAAGCTTTCGGGAAAGGAGTCCCTGAAATGCCAGATGTGGA 539
Oy 540 TGCCATCAACCAAGCTGCTGACTGCTGGGATATCCGCTGCTCCCTTATGATATCAAGA 599
Db 540 TGCCATCAACCAAGCTGCTGACTGCTGGGATATCCGCTGCTCCCTTATGATATCAAGA 599
Oy 600 TATCATGCTGCCACCCCGGGTGAAGAGTCTATGAGATGCAAGGTGAGCGCGAGCGCGG 659
Db 600 TATCATGCTGCCACCCCGGGTGAAGAGTCTATGAGATGCAAGGTGAGCGCGAGCGCGG 659
Oy 660 GAAACGGGCGCACAGTCTTGAAGTCTGAGGGGACCCGAGAGTCCGATCAATGTGGCAGA 719
Db 660 GAAACGGGCGCACAGTCTTGAAGTCTGAGGGGACCCGAGAGTCCGATCAATGTGGCAGA 719
Oy 720 AGGAGAGAAACAGGCGCCAGATCTGCTGCTCGAAGCAGAAAAGGCTGAACAGATAAATCA 779
Db 720 AGGAGAGAAACAGGCGCCAGATCTGCTGCTCGAAGCAGAAAAGGCTGAACAGATAAATCA 779
Oy 780 GCGACGACGAGAGCGGCACTGCAATTTCTGGCGAAGGCCAAGGCTAAAGCTAAGCTATTTCG 839
Db 780 GCGACGACGAGAGCGGCACTGCAATTTCTGGCGAAGGCCAAGGCTAAAGCTAAGCTATTTCG 839
Oy 840 AATCTGCTGCTCAGCTGTGACAAACATATGAGAGATGACAGCTTCACTGACGTGGGC 899
Db 840 AATCTGCTGCTCAGCTGTGACAAACATATGAGAGATGACAGCTTCACTGACGTGGGC 899
Oy 900 CGAGAGATATGTACAGCGGCTTCTCCAAACTGGGCCAAGGACTCAACACTATCTACTGCC 959
Db 900 CGAGAGATATGTACAGCGGCTTCTCCAAACTGGGCCAAGGACTCAACACTATCTACTGCC 959
Oy 960 CTCACACCTTGGCGATGTACACAGCATGCTGAGCTCAGGCTGTATTTGAGACCTT 1019
Db 960 CTCACACCTTGGCGATGTACACAGCATGCTGAGCTCAGGCTGTATTTGAGACCTT 1019
Oy 1020 CACCAAAAGCCCGAGTGCAGAGGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGT 1079
Db 1020 CACCAAAAGCCCGAGTGCAGAGGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGAGATGT 1079

```

```

Oy 1080 CCAGGCTACAGTCAAGTNTGATGAGAACTGATCGAGTCAAGATGATTAGTGAG 1139
Db 1080 CCAGGCTACAGTCAAGTNTGATGAGAACTGATCGAGTCAAGATGATTAGTGAG 1139
Oy 1140 CTGGGCTTNGCCAGGAGTCTGGGGACAAGAGAGATTTTCTGAT 1188
Db 1140 CTGGGCTTNGCCAGGAGTCTGGGGACAAGAGAGATTTTCTGAT 1188

```

RESULT 3

```

BC002442 1278 bp mRNA linear PRI 12-JUL-2001
LOCUS Homo sapiens, stomatin-like 2, clone MGC:1179 IMAGE:3346384, mRNA,
DEFINITION complete cds.
ACCESSION BC002442
VERSION BC002442.1 GI:12803254
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE 1 (bases 1 to 1278)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (05-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

```

```

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: qcqaps-remail.nih.gov
Tissue procurement: ATCC/DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 5 Row: m Column: 8
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 9652258.

Location/Qualifiers

1. 1278

/organism="Homo sapiens"

/db_xref="LocusID:30968"

/db_xref="taxon:9606"

/clone="MGC:1179 IMAGE:3346384"

/tissue_type="Skin, melanotic melanoma."

/clone_1lb="NIH_MGC_20"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

47. 1117

/codon_start=1

/product="stomatin-like 2"

/protein_id="AAH02442.1"

/db_xref="GI:12803255"

/translation="MLAARGTGALLRGLASGRAPRASSGLPRTNVLFPVPOG
 EAWVRMGRHRLIEPGLINTLIVLDRITRVOSKEIVINVPQSAVTLNATVLD
 GVLRIDMPYKASGVDEPREAYTOLAOTMRSELCISDKYPRRESINATVDA
 INQADWGRICLRLEIKDINHPRAVSKMOMVDAERRKRAVLESCTRESINVA
 EGRKQAOGLASEAKAEQINQAGEASVLAARAKAVALIILAAALDGHNGDAASL
 TVAEQVSAFSLAKDSNTILIPSNPDVTSVMVQAMGVGALTRKAPVPGPDSLSG

FEATURES

source

CDS

BASE COUNT 303 a 325 c 367 g 265 t
ORIGIN

Query Match 95.3%; Score 1132; DB 9; Length 1260;
Best Local Similarity 99.5%; Pred. No. 1.2e-258;
Matches 1154; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

30 TCGTGGTTCGAGAGTCTGCGGGGGGAAATGCGGGCGGGCGGGGCGA 89
1 TCGTGGTTCGAGAGTCTGCGGGGGGAAATGCGGGCGGGCGGGGCGA 59
90 CTGGGGCCCTTTTGGTGAAGGGCTCTACTGGCTTGTGGCCGCTCGG-3CCGCGCT 148
60 CTGGGGCCCTTTTGGTGAAGGGCTCTACTGGCTTGTGGCCGCTCGGCGCGGCT 119
149 CCTCTGGATTCGCGGAAACACCGTGTACTTGTGTGCGCGAGCAGAGGCTGGGTGG 208
120 CCTCTGGATTCGCGGAAACACCGTGTACTTGTGTGCGCGAGCAGAGGCTGGGTGG 179
209 TGGAGCGAATGGGCGCGATTCACCGCGATCTGGAGCGCTGGTTGAACATCCATCCCTG 268
180 TGGAGCGAATGGGCGCGATTCACCGCGATCTGGAGCGCTGGTTGAACATCCATCCCTG 239
269 TGTAGACCGGATCCGATATGTGACAGAGTCTCAGAGAAATTTGTCATCAACGTCCTGAGC 328
240 TGTAGACCGGATCCGATATGTGACAGAGTCTCAGAGAAATTTGTCATCAACGTCCTGAGC 299
329 AGTGGCTGTACTCTCGACATGTACTCTGCAATGATGATGAGTCTTTTACTGCGCA 388
300 AGTGGCTGTACTCTCGACATGTACTCTGCAATGATGATGAGTCTTTTACTGCGCA 359
389 TCATGACCGCTTACAGAGCAAGCTAGGAGTGTGAGAGCGCGTATGCGCGCACCGCAGC 448
360 TCATGACCGCTTACAGAGCAAGCTAGGAGTGTGAGAGCGCGTATGCGCGCACCGCAGC 419
449 TAGCTCAACAACATCATGATGAGTCTGCGCAACCTCTCTGTGACAAAGTCTTCCGGG 508
420 TAGCTCAACAACATCATGATGAGTCTGCGCAACCTCTCTGTGACAAAGTCTTCCGGG 479
509 AACGGAGTCCCTGATATCCAGCATTTGGATGCGCATACCAACAGCTGTGACTGCTGGG 568
480 AACGGAGTCCCTGATATCCAGCATTTGGATGCGCATACCAACAGCTGTGACTGCTGGG 539
569 GTATCGGCTGCTTGTATGATGATCAAGATATCATGATGCGCACCGCGGTGAAAGAGT 628
540 GTATCGGCTGCTTGTATGATGATCAAGATATCATGATGCGCACCGCGGTGAAAGAGT 599
629 CTATGACATGAGTGTGAGGAGCGAGCGGCGGAAACAGCTTCTAGAGTCTGAGG 688
600 CTATGACATGAGTGTGAGGAGCGAGCGGCGGAAACAGCTTCTAGAGTCTGAGG 659
689 GGACCCGAGATCGGCGCATCATGTGGCAGAAAGGAGGAGGCGGAGTCTGCGCT 748
660 GGACCCGAGATCGGCGCATCATGTGGCAGAAAGGAGGAGGCGGAGTCTGCGCT 719
749 CGAAGCGAAGAGCTGAAGCATTAATCAGGAGCAGGAGGCGGAGTCTGCGCT 808
720 CGAAGCGAAGAGCTGAAGCATTAATCAGGAGCAGGAGGCGGAGTCTGCGCT 779
809 CGAAGCGAAGAGCTGAAGCATTAATCAGGAGCAGGAGGCGGAGTCTGCGCT 868
780 CGAAGCGAAGAGCTGAAGCATTAATCAGGAGCAGGAGGCGGAGTCTGCGCT 839
869 ATGAGATGACAGAGCTTCACTGATGTGGCGGAGCAGTATGTCAAGCGCTTCTCAAC 928
840 ATGAGATGACAGAGCTTCACTGATGTGGCGGAGCAGTATGTCAAGCGCTTCTCAAC 899
929 TGGCGAAGAGCTCAACATCTACTGCGCTTCAACCGTGGGAGTGTCAACAGAGTGG 988
900 TGGCGAAGAGCTCAACATCTACTGCGCTTCAACCGTGGGAGTGTCAACAGAGTGG 959
989 TGGCTCAGAGCCATGGGTGTATATGAGGCGCTTCAACCAAGCGGAGTGTCAAG 1048

Db 960 TGGCTCAGAGCCATGGGTGTATATGAGCGCCCTCACCAAGCCCGAGCGGAGCTCCAG 1019
Qy 1049 ACTCACTCTCCAGTGGGAGCAGACAGATGTCCAGGATGACATGACATGTTGATGAG 1108
Db 1020 ACTCACTCTCCAGTGGGAGCAGACAGATGTCCAGGATGACATGACATGTTGATGAG 1079
Qy 1109 AACTTGATGACATCAAGATGATGATGAGTGTGAGGCTTGGCAGGAGTCTGGGAGCA 1168
Db 1080 AACTTGATGACATCAAGATGATGATGAGTGTGAGGCTTGGCAGGAGTCTGGGAGCA 1139
Qy 1169 GGAAGCAGATTTTCTGAT 1188
Db 1140 GGAAGCAGATTTTCTGAT 1159
RESULT 5
AK027405 1244 bp mRNA linear PRI 15-MAY-2001
LOCUS AK027405
DEFINITION Homo sapiens CDNA FLJ14499 fls, clone NT2RM1000080, weakly similar to UNC-1 PROTEIN.
ACCESSION AK027405
VERSION AK027405.1 GI:14042059
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens teratocarcinoma cell line:NT2 CDNA to mRNA, clone.lib:NT2RM1 clone:NT2RM1000080.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (sites)
Isigaki,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H., Magatsuma,M., Hosokiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K., Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murekami,K., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakematsu,A., Nakamura,Y., Nagahara,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T. NEBO human cDNA sequencing project Unpublished
2 (bases 1 to 1244)
Isigaki,T. and Otsuki,T.
REFERENCE 10-MAY-2001) Takao Isigaki, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp. Tel:81-438-52-3951. Fax:81-438-52-3952)
COMMENT NEBO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; CDNA library construction; Research Association for Biotechnology; CDNA library construction; 5' - 3' end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
FEATURES
source
Location/Qualifiers
1..1244
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RM1000080"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RM1"
/note="cloning vector: pUC19FL3-mRNA from uninduced NT2 neuronal precursor cells."
35..1105
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB55091.1"
/db_xref="GI:14042060"
/translation="MLARAGTGALLRGSLASGRAPRRASGLPNTVLFVPOO EAWVERGRHRHLEPGLNLIPLVDIRRVOSKEIVINWPDSATLIDNTYLOID GVLIRINDPKASTGVEDPEYATVOLAQTTRSLGLSLDKVFRRESLNASTVDA INQADWGIRCKLREYKIDHVPVPRVSKMOMOEAEERRKATVLESGTRESA INVA EGKROAOILASEARKEAOINOAGASASVILKAKAKAEALILAAALOHNGDAASL TVAEQYVAFSKLAKDSNTILLPSNPBDVTSMVAMQMVGYALTRKAPVPGPDSLSG

SSRDVQGTDSLDELDLRVKS"
 BASE COUNT 284 a 327 c 368 g 265 t
 ORIGIN

Query Match 95.28; Score 1130.4; DB 9; Length 1244;
 Best Local Similarity 99.48; Pred. No. 2, 9e-258;
 Matches 1153; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

30 TCGTGGTTCGGAGCGTGGCGGGCGGTGGGAAATGCTGGCGCGCGCGGGGCA 89
 1 TCGTGGTTCGGAGCGTGGCGGGCGGTGGGAAATGCTGGCGCGCGCGGGGCA 59
 90 CTGGGGCCCTTTTCTGCTGAGGGGCTCTACTGCGCTTGGCGCGCGCTCG- 148
 60 CTGGGGCCCTTTTCTGCTGAGGGGCTCTACTGCGCTTGGCGCGCGCGCGCT 119
 149 CCTCTGATTTGCCCGCAACACCGTGTACTGTTCTGTCGGCAGCAGAGAGCGCTGG 208
 120 CCTCTGATTTGCCCGCAACACCGTGTACTGTTCTGTCGGCAGCAGAGAGCGCTGG 179
 209 TGGAGCGAATGGGGCGATTCCACCGGATCCGAGCGCTGGTTTGAACATCCTCATCCCTG 268
 180 TGGAGCGAATGGGGCGATTCCACCGGATCCGAGCGCTGGTTTGAACATCCTCATCCCTG 239
 269 TGTTAGACCGGATCCGATATGTGAGAGTCTCAAGAAATTGTCAACGTCCTGAGC 328
 240 TGTTAGACCGGATCCGATATGTGAGAGTCTCAAGAAATTGTCAACGTCCTGAGC 299
 329 AGTGGCGTGTACTCTGAGCAATTAATCTGCAAAATGAGTGGAGTCTTAACTGGGCA 388
 300 AGTGGCGTGTACTCTGAGCAATTAATCTGCAAAATGAGTGGAGTCTTAACTGGGCA 359
 389 TCATGAGACCCCTTACAGAGCAAGTACGGTGTGAGAGACCCCTGAGTATGCCCTCAACCAGC 448
 360 TCATGAGACCCCTTACAGAGCAAGTACGGTGTGAGAGACCCCTGAGTATGCCCTCAACCAGC 419
 449 TAGCTCAAAACACCATGATGATCAGAGCTCGGCAAACTCTCTNTGGACAAATCTTTCGGGG 508
 420 TAGCTCAAAACACCATGATGATCAGAGCTCGGCAAACTCTCTNTGGACAAATCTTTCGGGG 479
 509 AAGGGAGTCCCTGAATGCCAGCATTTGGATGCCATCAACCAAGCTCTGACTGCTGG 568
 480 AAGGGAGTCCCTGAATGCCAGCATTTGGATGCCATCAACCAAGCTCTGACTGCTGG 539
 569 GTATCCGCTGCCCTTCGTTATGATCAAGATATCCATGTCCACCCCGGGTGAAGAGT 628
 540 GTATCCGCTGCCCTTCGTTATGATCAAGATATCCATGTCCACCCCGGGTGAAGAGT 599
 629 CTATGACAGATGAGTGGAGCAGAGCGGGGAAACGGGCCACAGTTCTAGAGTCTGAGG 688
 600 CTATGACAGATGAGTGGAGCAGAGCGGGGAAACGGGCCACAGTTCTAGAGTCTGAGG 659
 689 GGACCCGAGATCGGCGCATCATATGTGAGAGAGAGAGAAACAGAGCCAGATCCTGGCCT 748
 660 GGACCCGAGATCGGCGCATCATATGTGAGAGAGAGAGAAACAGAGCCAGATCCTGGCCT 719
 749 CCGAAGCAGAAAAGGCTGACAGATAAATCAGAGCAGCAGAGAGGCCAGTGCAGTTCTGG 808
 720 CCGAAGCAGAAAAGGCTGACAGATAAATCAGAGCAGCAGAGAGGCCAGTGCAGTTCTGG 779
 809 CGAAGGCCAAGGCTAAAGCTAAGCTATTGCAATCTGGCTGACCTGACACAAACAT 868
 780 CGAAGGCCAAGGCTAAAGCTAAGCTATTGCAATCTGGCTGACCTGACACAAACAT 839
 869 ATGAGATGACAGACACTTCACTGATGTGGCCGAGACAGATATGTACAGCGTTCTCCAAAC 928
 840 ATGAGATGACAGACACTTCACTGATGTGGCCGAGACAGATATGTACAGCGTTCTCCAAAC 899
 929 TGGCCAAAGACTCAACACTATCTACTGCGCTTCCAAACCTGGCGCATGTCAACAGCATGG 988
 900 TGGCCAAAGACTCAACACTATCTACTGCGCTTCCAAACCTGGCGCATGTCAACAGCATGG 959

QY 989 TGGCTAGGCCATGGGTGTATATGAGACCCCTCAACAAAGCCCGAGTGGAGCTCAG 1048
 DB 960 TGGCTAGGCCATGGGTGTATATGAGACCCCTCAACAAAGCCCGAGTGGAGCTCAG 1019
 QY 1049 ACTACACTCCAGTGGAGAGCAGAGATGTCCAGGGGTACAGATGCAAGTNTGTATGAGG 1108
 DB 1020 ACTACACTCCAGTGGAGAGCAGAGATGTCCAGGGGTACAGATGCAAGTNTGTATGAGG 1079
 QY 1109 AACTTGATGAGTCAAGATGATTAGTGGAGCTGGGCTTGGCCAGGAGTCTGGGCAAA 1168
 DB 1080 AACTTGATGAGTCAAGATGATTAGTGGAGCTGGGCTTGGCCAGGAGTCTGGGCAAA 1139
 QY 1169 GGAAGCAGATTTTCTGATT 1188
 DB 1140 GGAAGCAGATTTTCTGATT 1159

RESULT 6

BC010152 1281 bp mRNA 11near PRI 12-JUL-2001
 LOCUS
 DEFINITION Homo sapiens, stomatin-like 2, clone MGC:19715 IMAGE:3534656, mRNA,
 complete cds.
 ACCESSION BC010152
 VERSION BC010152.1 GI:14603402
 KEYWORDS MGC.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1281)
 Strausberg, R.
 Direct Submission
 Submitted (02-JUL-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REMARK
 COMMENT
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: gcgabs-remail.nih.gov
 Tissue Procurement: DCTD/DTP
 cDNA Library Preparation: Rudin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web Site: <http://www.nisc.nih.gov/>
 Contact: nisc.mgc@nih.gov
 Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
 Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
 Lim, M., Maduro, O.L., Mastello, C., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Snyder, B., Stentrop, S., Thomas, P.J.,
 Tjongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, D.L., Walker, M.A.,
 Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 27 Row: d Column: 1
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Similarity but not
 identity to protein.

FEATURES
 source

1..1281
 /organism="Homo sapiens"
 /db_xref="LocusID:30968"
 /db_xref="taxon:9606"
 /clone="MGC:19715 IMAGE:3534656"
 /tissue_type="Lung, small cell carcinoma"
 /clone_lib="NHL MGC-7"
 /lab_host="DH10B-R"
 /note="Vector: pOT87"
 44..1114
 /codon_start=1

CDS

```

/product="stomatin-like 2"
/protein_id="AAH10152.1"
/db_xref="gi:14603403"
/translation="MLAARAGCTGALLIRGLSLASGRAPRRASGLPENTVLEVPPO
EAMVBERMGRFRHRLERPLNLIIVLYDRIRYVSLKEIVINPPOSATVLTNATLQID
GVLTLRIMDPYKASYGVEDPEYAVTQPRQTMRSLEGLSLDKFRESRESINASIVDA
INQADWGIJRCLEKYEIKDIDHPRVKESLMQKQVLEAEKREKATVLESGTRESINVA
EGKQAOIILASEAKAEQIINOAGEASVILAKAKAEKAIILAAALTOHNGDAASL
TVAEQYVAFSKLAKDSNTLILPSNPQVTSMAVQAGVYALIRKAPVPGTPDLSLSSG
SSRDVQGTMDASLDELDRVKMS"
BASE COUNT      315 a      330 c      372 g      264 t
ORIGIN

```

```

Query Match      95.2%; Score 1130.4; DB 9; Length 1281;
Best Local Similarity 99.4%; Pred. No. 2.9e-258;
Matches 1153; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

QY 30 TCGTTGGTTCCGGAGGTCGTCGGCGGGTGGGAAATGCTGGCGCGCGCGGGGGA 89
    |||||||
Db 10 TCGTTGGTTCCGGAGGTCGTCGGCGGGTGGGAAATGCTGGCGCGCGCGGGGGA 68
    |||||||

QY 90 CTGGGGCCCTTTTGGTGAGGGGCTCTTACTGGCTTTGGCCGGCTCG -GCCGCGCT 148
    |||||||
Db 69 CTGGGGCCCTTTTGGTGAGGGGCTCTTACTGGCTTTGGCCGGCTCGCGCGCGCT 128
    |||||||

QY 149 CCTGTGATGCCCCGGAACACCGGTACTGTCTGCGCCGACGAGAGGCTGGGTG 208
    |||||||
Db 129 CCTGTGATGCCCCGGAACACCGGTACTGTCTGCGCCGACGAGAGGCTGGGTG 188
    |||||||

QY 209 TGGACGGAATGGCGGATTCACACCGGATCTGGAGCTGTGTTGAACATCCATCCCTG 268
    |||||||
Db 189 TGGACGGAATGGCGGATTCACACCGGATCTGGAGCTGTGTTGAACATCCATCCCTG 248
    |||||||

QY 269 TGTTAGACCGGATCCGATATGTGCAGACTCTCAAGAAATTTCTCAATCAAGCTGAGC 328
    |||||||
Db 249 TGTTAGACCGGATCCGATATGTGCAGACTCTCAAGAAATTTCTCAATCAAGCTGAGC 308
    |||||||

QY 329 AGTCGGCTGTGACTCTGCACATGTAATCTGCAATGAGATGAGATCCCTTACCTGGCGCA 388
    |||||||
Db 309 AGTCGGCTGTGACTCTGCACATGTAATCTGCAATGAGATGAGATCCCTTACCTGGCGCA 368
    |||||||

QY 389 TCATGGACCCCTTACAAAGGCAAGCTAGGTGTGGAGAGACCCCTGATGATGCCGTACCCAGC 448
    |||||||
Db 369 TCATGGACCCCTTACAAAGGCAAGCTAGGTGTGGAGAGACCCCTGATGATGCCGTACCCAGC 428
    |||||||

QY 449 TAGCTCAAAACAACCTAGATCAGAGCTCGGCAAACTCTCTTTGGAACAAAGTCTTCGCGG 508
    |||||||
Db 429 CAGCTCAAAACAACCTAGATCAGAGCTCGGCAAACTCTCTCTGGAACAAAGTCTTCGCGG 488
    |||||||

QY 509 AACGGAGATCCCTGAAATGCAGCATTTGGAGTGCATCAACCAAGCTGTAATCTGCTGGG 568
    |||||||
Db 489 AACGGAGATCCCTGAAATGCAGCATTTGGAGTGCATCAACCAAGCTGTAATCTGCTGGG 548
    |||||||

QY 569 GTATCCGCTGCTTNCCTTATGATCAAGATATCATGTCACCCCGG3TGAAGAGT 628
    |||||||
Db 549 GTATCCGCTGCTTNCCTTATGATCAAGATATCATGTCACCCCGG3TGAAGAGT 608
    |||||||

QY 629 CTATGACATGAGGTGAGAGGAGCGCGGGAACGGGCCACAGTTCTAGAGTCTGAGG 688
    |||||||
Db 609 CTATGACATGAGGTGAGAGGAGCGCGGGAACGGGCCACAGTTCTAGAGTCTGAGG 668
    |||||||

QY 689 GGACCCGAGAGTGGCCCATATGTCAGAGGGAAGGAACAGGCCCAAGATCCGCGCT 748
    |||||||
Db 669 GGACCCGAGAGTGGCCCATATGTCAGAGGGAAGGAACAGGCCCAAGATCCGCGCT 728
    |||||||

QY 749 CCGAAGCAGAAAAAGCTGACAGATTAATCAAGCAGACAGAGAGGCCAGTGCAGTTCTGG 808
    |||||||
Db 729 CCGAAGCAGAAAAAGCTGACAGATTAATCAAGCAGACAGAGAGGCCAGTGCAGTTCTGG 788
    |||||||

QY 809 CGAAGGCCAAGGCTAAAGCTGAAGCTATTTCATCTGGCTGACGCTGACACAACATA 868
    |||||||
Db 789 CGAAGGCCAAGGCTAAAGCTGAAGCTATTTCATCTGGCTGACGCTGACACAACATA 848
    |||||||

```

```

QY 869 ATGAGATGCACAGCTTTCAGTACTGATGCGCGGAGAGATGTGACGGGCTTCCAAAC 928
    |||||||
Db 849 ATGAGATGCACAGCTTTCAGTACTGATGCGCGGAGAGATGTGACGGGCTTCCAAAC 908
    |||||||

QY 929 TGGCCAAAGAGCTCAACACTATATCTACTGCTCCCAACCCCTGGCATGTCCACAGCATGG 988
    |||||||
Db 909 TGGCCAAAGAGCTCAACACTATATCTACTGCTCCCAACCCCTGGCATGTCCACAGCATGG 968
    |||||||

QY 989 TGGCTCAGAGCCATGGGTGTATGTGAGCCCTTACCAAAAGCCCAAGTGCAGGACTCCAG 1048
    |||||||
Db 969 TGGCTCAGAGCCATGGGTGTATGTGAGCCCTTACCAAAAGCCCAAGTGCAGGACTCCAG 1028
    |||||||

QY 1049 ACTCACTCTCCAGTGGAGACAGAGATGTCCAGAGTACAGATGCAAGTNTTGATGAG 1108
    |||||||
Db 1029 ACTCACTCTCCAGTGGAGACAGAGATGTCCAGAGTACAGATGCAAGTNTTGATGAG 1088
    |||||||

QY 1109 AACTGTATGAGTCAAGATGATGAGTGGGCTTGGCCAGGAGTCTGGGGACAA 1168
    |||||||
Db 1089 AACTGTATGAGTCAAGATGATGAGTGGGCTTGGCCAGGAGTCTGGGGACAA 1148
    |||||||

QY 1169 GGAAGCAGATTTTCTGATT 1188
    |||||||
Db 1149 GGAAGCAGATTTTCTGATT 1168
    |||||||

```

RESULT 7

```

BC014990
LOCUS      BC014990                1262 bp      mRNA           linear      PRI 29-OCT-2001
DEFINITION Homo sapiens, stomatin-like 2, clone MGC:23225 IMAGE:4907745, mRNA, complete cds.
ACCESSION  BC014990
VERSION    BC014990.1  GI:15929069
KEYWORDS   MGC.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE  1 (bases 1 to 1262)
AUTHORS   Strausberg, R.
TITLE     Direct Submission
JOURNAL    Submitted (01-OCT-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA

```

```

REMARK     NIH-MGC Project URL: http://mgc.nci.nih.gov
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC

```

```

            CDNA Library Preparation: Rubin Laboratory
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca

```

```

            Steven Jones, Jennifer Asano, Ian Bosdel, Yaron Butterfield,
            Susana Chan, Readman Chiu, Chris Fell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kuesche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Ilisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smalins, Michael Smith, Lorraine Spence, Jeff Stolt,
            Michael Thorne, Miranada Teal, Nastasia van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Maira.

```

```

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 3-E. Consortium/LLNL at: http://image.llnl.gov
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Similarity but not
            identity to protein

```

```

FEATURES
            location/Qualifiers
            1..1262

```

```

            /organism="Homo sapiens"
            /db_xref="LocusID:30968"

```

```

/db_xref="taxon:9606"
/clone="MGC:21225 IMAGE:4907745"
/tissue="Pancreas; epithelioid carcinoma"
/clone_lib="NIH.MGC_42"
/lab_host="DH10B-R"
/notes="Vector: pOT87"
33..1103
/codon_start=1
/product="stomatin-like 2"
/protein_id="AAH14990.1"
/db_xref="GI:15929070"
/translation="MLAARAGTGAALLRSLASGRAPRRASGLPRNTVYLFPOQ
EAWVERMGRFHRILEPGLNLIPLVDLRIRVOSLEKIVINPEQSAVTLDNVLOID
GVYLRIMDPYKASVGEDEYAVTOLAOTRMSLEKRLSLDKYFERESIASIVA
INQADWGIKRLYEIKIDIHVPKRESHQMYEARRRKATVLESEGTESAIYA
EKGKAOIILASEAKAEIOINQAGASAVILAKAKAEATRIILAAVLOHNGDAASL
TVAEQVYSAFSKLAKDSNTLLIPSPNDVYSNVAQAMGVIGALKIAPGPDPSSLSSG
SSRDVQCTDASLDELDKRVKMS"
BASE COUNT      305 a      326 c      367 g      264 t
ORIGIN
Query Match      95.1%   Score 1130:   DB 9:   Length 1262:
Best Local Similarity 99.5%   Pred. No. 3.6e-258:
Matches 1152:   Conservative 0:   Mismatches 4:   Indels 2:   Gaps 2:
QY  32 GTTGGTTCGGAGAGTCGTGGCGGCGGTGGAATGCTGGCGCGCGGCGGGGCACT 91
Db   1 GTTGGTTCGGAGAGTCGTGGCGGCGGTGGAATGCTGGCGCGCGGCGGGCACT 59
QY  92 GGGGCGCTTTTGTCTGAGAGGCGCTCTACTGGCTTGTGGCGCGGCTCCG-GCCGCGCTTC 150
Db   60 GGGGCGCTTTTGTCTGAGAGGCGCTCTACTGGCTTGTGGCGCGGCTCCGCGCGCTTC 119
QY  151 TCTGGATTGCCCCGGAACACCGGTGACTGTTGTCGCCGACAGAGAGGCGGTGGTG 210
Db  120 TCTGGATTGCCCCGGAACACCGGTGACTGTTGTCGCCGACAGAGAGGCGGTGGTG 179
QY  211 GAGGGAATGGCGCGGATTCACCGGATCCTGGAGCGCTGGTTGAACATCCTATCCCTGTG 270
Db  180 GAGGGAATGGCGCGGATTCACCGGATCCTGGAGCGCTGGTTGAACATCCTATCCCTGTG 239
QY  271 TTAAACCGGATTCCTATGTCGAGAGTCTCAAGGAATTTGTCATCAACAGTGGCTAGCAG 330
Db  240 TTAAACCGGATTCCTATGTCGAGAGTCTCAAGGAATTTGTCATCAACAGTGGCTAGCAG 299
QY  331 TCGGCTGTGACTTCGACATGTAACCTGCAAAATCGATGAGTCTTACCTGGCAATC 390
Db  300 TCGGCTGTGACTTCGACATGTAACCTGCAAAATCGATGAGTCTTACCTGGCAATC 359
QY  391 ATGACACCTTACAGAGCAAGCTACGGTGTGAGAGACCCCTGATGAGTACCCAGCCTA 450
Db  360 ATGACACCTTACAGAGCAAGCTACGGTGTGAGAGACCCCTGATGAGTACCCAGCCTA 419
QY  451 GCTAAACCAACCATGAGATTCAGAGTCTGGCAAACTCTTNTGGAACAAAGTCTTCCGGGAA 510
Db  420 GCTAAACCAACCATGAGATTCAGAGTCTGGCAAACTCTTNTGGAACAAAGTCTTCCGGGAA 479
QY  511 CGGAGAGTCCCTGAATGCCAGATTTGTGATGCCATCAACCAAGCTGTGAGTCTGGGGT 570
Db  480 CGGAGAGTCCCTGAATGCCAGATTTGTGATGCCATCAACCAAGCTGTGAGTCTGGGGT 539
QY  571 ATCCGCTCCCTNCGTTATGAGATCAAGATATCCATGTGCCACCCGGGTGAAGAAGTCT 630
Db  540 ATCCGCTCCCTNCGTTATGAGATCAAGATATCCATGTGCCACCCGGGTGAAGAAGTCT 599
QY  631 ATGCAATGCAAGTGTGAGAGGAGCGGCGGAAACGGGCGACAGTTCTAGAGTCTAGAGGG 690
Db  600 ATGCAATGCAAGTGTGAGAGGAGCGGCGGAAACGGGCGACAGTTCTAGAGTCTAGAGGG 659
QY  691 ACCGAGAGTGGCGCATCAATGTGGCAGAGGAAGAAACAGGCCACAGATCCTGGCCTCC 750
Db  660 ACCGAGAGTGGCGCATCAATGTGGCAGAGGAAGAAACAGGCCACAGATCCTGGCCTCC 719

```

```

QY  751 GAAGCAGAAAAGGCTGAACAGATAAATCAGCAGCAGAGAGAGGCCAGTGTCTGGCG 810
Db  720 GAAGCAGAAAAGGCTGAACAGATAAATCAGCAGCAGAGAGAGGCCAGTGTCTGGCG 779
QY  811 AAGGCCAAGGCTAAAGCTGAAGCTATTGTAATCTGTGCTGCAGCTCTGACACCAATAT 870
Db  780 AAGGCCAAGGCTAAAGCTGAAGCTATTGTAATCTGTGCTGCAGCTCTGACACCAATAT 839
QY  871 GGAGATGAGAGAGCTTCACTGACTGTGGCGCAGAGATGTGTCAGGCGGCTTCCAAACTG 930
Db  840 GGAGATGAGAGAGCTTCACTGACTGTGGCGCAGAGATGTGTCAGGCGGCTTCCAAACTG 899
QY  931 GCCAAGAGCTCCACACATATCTACTGCGCTCCAAACCTGGCGATGTACACAGCATGTG 990
Db  900 GCCAAGAGCTCCACACATATCTACTGCGCTCCAAACCTGGCGATGTACACAGCATGTG 959
QY  991 GCTACGCGCATGTGTATATGAGAGCCTCAACCAAGCCCAAGTCCAGAGCTCCAGAC 1050
Db  960 GCTACGCGCATGTGTATATGAGAGCCTCAACCAAGCCCAAGTCCAGAGCTCCAGAC 1019
QY  1051 TCACCTCCAGAGGAGGAGCAGAGATGTCCAGGCTACAGATGCAAGNTTGATGAGGAA 1110
Db  1020 TCACCTCCAGAGGAGGAGCAGAGATGTCCAGGCTACAGATGCAAGNTTGATGAGGAA 1079
QY  1111 CTTGATCGAGTCAAGATAGTTAGTGAGAGCTGGGCTTNGCCAGGAGAGTCTGGGACAAAG 1170
Db  1080 CTTGATCGAGTCAAGATAGTTAGTGAGAGCTGGGCTTNGCCAGGAGAGTCTGGGACAAAG 1139
QY  1171 AAGCAGATTTTCTCTGATT 1188
Db  1140 AAGCAGATTTTCTCTGATT 1157

```

```

RESULT 8
LOCUS BC003025 1257 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, stomatin-like 2, clone MGC:4191 IMAGE:2821269, mRNA,
complete cds.
ACCESSION BC003025.
VERSION BC003025.1 GI:12804332
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1257)
REFERENCE
1 (bases 1 to 1257)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
NIH-MGC Project URL: http://mgc.nci.nih.gov
CONTACT: MGC help desk
Email: gcaps-remail.nih.gov
Tissue Procurement: DCRD/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
Info@cgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhui, Parvaneh Saeedi, Jacqueline
Schelin, Duane Smallus, Michael Smith, Lorraine Spence, Jeff Stolt,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 9 Row: 1 Column: 7
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 9652258.
Location/Qualifiers

FEATURES

source

1..1257
/organism="Homo sapiens"
/db_xref="taxon:9606"
/db_xref="locusID:30968"
/clone="MGC:4191 IMAGE:2821269"
/tissue_type="Lung, small cell carcinoma"
/clone_id="NH_MGC_7"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
23..1093
/codon_start=1
/product="stomatin-like 2"
/protein_id="AA03025.1"
/db_xref="GI:12804333"
/translation="MLAAARGTGAALLRGSILASGRAPRRASGLPNTVVLFPVPO
EAMVVERMGRFRHLEPGLNLIPLVDLRIRYVOSLKEIVINPEOSAVTLNVTLOID
GLYLIRIDPYKASVYEDPEYAVTOLAOTMRSLEKLSIDKYERRESINASTVDA
INOADCMGTRCLREIKDIIHPPRVKSMOYEAERKRAIVLESSTRESATNVA
EGKQAOILASEAKAEKEDINQAGEASVILAKAKABATITILAAALTOHNDAAASL
TVABQVSAFSLAKDSNTILPSPNPGDVTSMVAQAMGVATLAKAPVGPDPSSLSG
SSRDVQGDALDELDRVKMS"

CDS

BASE COUNT 309 a 324 c 364 g 260 t
ORIGIN

Query Match 94.3%; Score 1120; DB 9; Length 1257;
Best Local Similarity 99.5%; Pred. No. 8.5e-256;
Matches 1142; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

OY 42 GAGGTCGCTGGGGGGTGGGAATGCTGGCGCGCGCGGCGGCACTGGGCGCTTT 101
DB 1 GAGGTCGCTGGGGGGTGGGAATGCTGGCGCGCGCGGCGGCGGCACTGGGCGCTTT 59
OY 102 TGCCTGAGGGGCTCTCTACTGCTTGGCGCGGCTCCG-GCGCGCGCTCTTGATTCG 160
DB 60 TGCCTGAGGGGCTCTCTACTGCTTGGCGCGGCTCCGCGCGCGGCGGCGCTCTTGATTCG 119
OY 161 CCCGAACACCGTGTACTGTTGCTGCGCGACAGAGAGGCGCTGGGTGTGAAGCAATGG 220
DB 120 CCCGAACACCGTGTACTGTTGCTGCGCGACAGAGAGGCGCTGGGTGTGAAGCAATGG 179
OY 221 GCCGATTCACCGGATCCTGGAGCCTGGTTGAACATCCTCTACCTGGGCATCAAGACCTTT 280
DB 180 GCCGATTCACCGGATCCTGGAGCCTGGTTGAACATCCTCTACCTGGGCATCAAGACCTTT 239
OY 281 TCCGATATGTCAGAGATCTCAAGAAATGTCATCAACGTGCTGAGAGTGGCTGTGA 340
DB 240 TCCGATATGTCAGAGATCTCAAGAAATGTCATCAACGTGCTGAGAGTGGCTGTGA 299
OY 341 CTCTGCAATGTAACTCTGCAAAATCGATGAGTCCCTTACTGCGCATCAAGACCTTT 400
DB 300 CTCTGCAATGTAACTCTGCAAAATCGATGAGTCCCTTACTGCGCATCAAGACCTTT 359
OY 401 ACAAGGCAAGCTAGGCTGGAGACCTGAGTATGCCCTACCCAGCTAGTCAACAA 460
DB 360 ACAAGGCAAGCTAGGCTGGAGACCTGAGTATGCCCTACCCAGCTAGTCAACAA 419
OY 461 CCATGAGATCAGAGCTCGGCAAACTCTCTNTGACAAAGTCTTCGGGGAAGGAGATCCC 520
DB 420 CCATGAGATCAGAGCTCGGCAAACTCTCTNTGACAAAGTCTTCGGGGAAGGAGATCCC 479
OY 521 TGAATGCCAGCACTTGTGATGCCATCAACAAAGCTGCTGAGTGGGGTATCCGCTGCC 580
DB 480 TGAATGCCAGCACTTGTGATGCCATCAACAAAGCTGCTGAGTGGGGTATCCGCTGCC 539
OY 581 TNCGTATGAGATCAAGGATATCCATGTGCCAAGCCCGGGTGAAGAGTATGACAGATGC 640
DB 540 TCCGTATGAGATCAAGGATATCCATGTGCCAAGCCCGGGTGAAGAGTATGACAGATGC 599

OY 641 AGGTGAGGCGAGAGCGCGGAAACGGGCCACAGTTCATAGTCTGAGGGGACCCGAGACT 700
DB 600 AGGTGAGGCGAGAGCGCGGAAACGGGCCACAGTTCATAGTCTGAGGGGACCCGAGACT 659
OY 701 CGGCCATCAATGTGGCAGAAAGGAAAGAAACAGGCCAGATCTTGCCCTCCGAAGCAAA 760
DB 660 CGGCCATCAATGTGGCAGAAAGGAAAGAAACAGGCCAGATCTTGCCCTCCGAAGCAAA 719
OY 761 AGGTGAGGCAAGTAAATTCAGGCGAGAGAGAGGCCAGTTCAGTTCGGGGAAGGCCAAG 820
DB 720 AGGTGAGGCAAGTAAATTCAGGCGAGAGAGAGGCCAGTTCAGTTCGGGGAAGGCCAAG 779
OY 821 CTAAAGCTGAAGCTATTCGAATCTGGCTGCAGCTCTGACACAAACATTAATGAGATGAC 880
DB 780 CTAAAGCTGAAGCTATTCGAATCTGGCTGCAGCTCTGACACAAACATTAATGAGATGAC 839
OY 881 CAGCTTCACTGACTGTGGCGGAGACAGTATGTCAGCGGCTTTCGCAAACTGGCCAGAGCT 940
DB 840 CAGCTTCACTGACTGTGGCGGAGAGAGTATGTCAGCGGCTTTCGCAAACTGGCCAGAGCT 899
OY 941 CCAACACTATCTTACTAGCCCTCCAAACCCCTGGGAGTGTACACAGCAATGTCAGAGCCA 1000
DB 900 CCAACACTATCTTACTAGCCCTCCAAACCCCTGGGAGTGTACACAGCAATGTCAGAGCCA 959
OY 1001 TGGGTGTATATGAGAGCCCTCACCAAGCCCAAGTGCAGAGACTCCAGACTCACTTCCA 1060
DB 960 TGGGTGTATATGAGAGCCCTCACCAAGCCCAAGTGCAGAGACTCCAGACTCACTTCCA 1019
OY 1061 GTGGGAGCAGCAGAGATGTCACAGGCTACAGATGCAAGTMTGATGAGGAATTTGATCGAG 1120
DB 1020 GTGGGAGCAGCAGAGATGTCACAGGCTACAGATGCAAGTMTGATGAGGAATTTGATCGAG 1079
OY 1121 TCAAGATGAGTATGAGAGACTGGGCTTCCAGCGGAGCTCGGGGAAAGCAACATTT 1180
DB 1080 TCAAGATGAGTATGAGAGACTGGGCTTCCAGCGGAGCTCGGGGAAAGCAACATTT 1139
OY 1181 TCCGTGATT 1188
DB 1140 TCCGTGATT 1147

RESULT 9
AF323178 1253 bp mRNA linear ROD 23-JAN-2001
LOCUS AF323178
DEFINITION Mus musculus stomatin-like protein 2 mRNA, complete cds.
ACCESSION AF323178
VERSION AF323178.1 GI:12382776
KEYWORDS
SOURCE
ORGANISM
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 1253)
AUTHORS Chang, J.G. and Chan, W.L.
TITLE Mouse stomatin-like protein 2 (MSLP2) mRNA
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 1253)
AUTHORS Chang, J.G. and Chan, W.L.
TITLE Direct Submission
JOURNAL Submitted (21-NOV-2000) Department of Medical Research, China
Medical College Hospital, 2, Yuh-Der Road, Taichung 404, Taiwan
FEATURES
source
1..1253
/organism="Mus musculus"
/strain="FVB"
/db_xref="taxon:10090"
/tissue_type="kidney"
32..1093
/note="MSLP2"
/codon_start=1
/product="stomatin-like protein 2"
/protein_id="AAG53404.1"
/db_xref="GI:12382777"

```

/translaction="MLARAARGTALLRSGVQAGRVPRASSGLPNTVILFVPOO
EAMVVRGRHRIIEPLNLIPLDIRIVOSLKEIVINVPQASATLNDVTLQID
GLVLRIMDPYKASYGVEDPEAVTOLAQTRISELSDKFEKRESESNANIVDA
INOMADCMGIRLEIKDIDHPPRVESOMOVIEBKRPRATVLESGTESAIYVA
EGKMOAOIILASEAKAEIOINQAGEASAVLAKAKAEARFILLAGATLOHNGDAASL
TVABOYVSABSKLAKDSNTVLLPSNPDSVTSMAVOMGVIALTKAPVGAONSSOR
RDVATDTSTLEELRVKLIS"
BASE COUNT      320 a      311 c      364 g      258 t
ORIGIN

```

```

Query Match      73.0%: Score 867.4; DB 10; Length 1253;
Best Local Similarity 88.0%: Pred. No. 9.7e-196;
Matches 1002; Conservative 0; Mismatches 125; Indels 12; Gaps 5;

QY 34 TGGTTCGGAGGCTGCTGCGGGTGGGAAATGTCGGCGCGCGGGGGGCACTGG 93
DB 2 TGGTTCGGAGGCTGCGGGTGGGAAATGTCGGCGCGGGGGGCACTGG 60

QY 94 GGGCTTTTGTGAGAGGCTCTCTACTGCTTGTGGCGCGCTCG-6CGCGCTCTC 152
DB 61 AGCCCTTTTGTGAGAGGCTCTCTACTGCTTGTGGCGCGCTCGCGCGCTCTC 120

QY 153 TGGATTGCGCGGAACACCGTGTCTGCTGCGCGAGAGAGCGCTGGGTGGTGA 212
DB 121 TGGATTGCGCGGAACACCGTGTCTGCTGCGCGAGAGAGCGCTGGGTGGTGA 180

QY 213 GCGAATGGGCGGATTCACACCGGATCTGGAGCGCTGTTGAACATCATCTGTGT 272
DB 181 GCGAATGGGCGGATTCACACCGGATCTGGAGCGCTGTTGAACATCATCTGTGT 240

QY 273 AGACCGGATTCGATATGTCAGAGTCTCAAGAAATGTTCATCAACGTGCGAGCAGTC 332
DB 241 AGACCGGATTCGATATGTCAGAGTCTCAAGAAATGTTCATCAACGTGCGAGCAGTC 300

QY 333 GGGTGTGCTCTCGCATGTAACTCTGCAATTCATGAGTCCCTTACCTCGCATCAT 392
DB 301 AGCGCTTACTCTTGCAATGTAACTCTGCAATTCATGAGTCCCTTACCTCGCATCAT 360

QY 393 GGACCTTTTACAGGCAAGCTACGCTGTGGAGAGCGCTGATGTCGCGCACCCAGCAGTCAGC 452
DB 361 GGATTCCTTACAGGCAAGCTACGCTGTGGAGAGCGCTGATGTCGCGCACCCAGCAGTCAGC 420

QY 453 TCACCAACATCATGATGATGAGAGCTGCGCAACTCTCTTGTGACAAAGTCTTCCGGAGAG 512
DB 421 TCAGAGCTCTGATGATGATGAGAGCTTGGCAACTCTCTGACAAAGTCTTCCGGAGAG 480

QY 513 GGAGTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 572
DB 481 TGAGTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540

QY 573 CGCGTCCCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 632
DB 541 CGGATGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600

QY 633 GCAGATGAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 692
DB 601 GCAGATGAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660

QY 693 CCGAGTGGGCGCATCATGTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 752
DB 661 AGCAGAGTACCTATTTATGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720

QY 753 AGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 812
DB 721 AGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780

QY 813 GGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 872
DB 781 GGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840

QY 873 AGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 932
DB 840 AGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900

```

```

DB 841 AGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 933 CAAGAGCTCAACACTATTCATCTGCTCCCTCCCAACCTGGAGATGACAGAGAGAGAGAGAGAGAGAGAG
DB 901 CAAGAGCTCAACACTATTCATCTGCTCCCTCCCAACCTGGAGATGACAGAGAGAGAGAGAGAGAGAGAG
QY 993 TCAGGCGATGGGTGTATATGAGAGCGCTTCACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 961 TCAGGCGATGGGTGTATATGAGAGCGCTTCACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 1053 ACTTCAGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 1015 GAGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 1113 TGGATGAGTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
DB 1072 GGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT

RESULT 10
BC003425
LOCUS
DEFINITION
MUS musculus, RIKEN cDNA 0610038F01 gene, clone MGC:6777
IMAGE:2609584, mRNA, complete cds.
ACCESSION
BC003425
VERSION
BC003425.1 GI:13097353
KEYWORDS
MGC.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 1565)
AUTHORS
Strausberg, R.
TITLE
Direct Submission
JOURNALS
Submitted (20-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: mgc@ncl.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (HLNI)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: Villalobos, tmc.edu.
Villalobos, D.K., Luna, R.A., Hale, S.M., Huliy, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.
FEATURES
source
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/HLNI at: http://image.llnl.gov
Series: IRAC Plate: 5 Row: K Column: 5.
Location/Qualifiers
1..1565
/organism="Mus musculus"
/db_xref="LocusID:66592"
/db_xref="taxon:10090"
/clone="MGC:6777 IMAGE:2609584"
/tissue-type="Mammary tumor, C3(1)-Tag model. Infiltrating
ductal carcinoma. 5 month old virgin mouse."
/clone_lib="NCL CGAP_Mam6"
/lab_host="DH10B"
/Note="Vector: pCMV-SPORT6"
12..1073
/codon_start=1
/product="RIKEN cDNA 0610038F01 gene"
/protein_id="AAH03425.1"
/db_xref="GI:13097353"
/translaction="MLARAARGTALLRSGVQAGRVPRASSGLPNTVILFVPOO

```


[illegible]

KEYWORDS	HG: HNGS_PHASE1.
ORGANISM	Norway rat.
ONCINIS	Rattus norvegicus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
AUTHORS	1 (bases 1 to 186656) Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alstbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbara,J., Benton,J., Blinage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burnett,C., Burrell,K.L., Byrd,N.C., Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Donathale,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frintz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B., Homsji,F., Howard,S., Huber,J., Hulys,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudan,S., Karlsson,E., Kelly,S., Khan,O., King,L., Korvah,J., Kovach,C., Kratovic,J., Kuresh,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,D., Li,Z., Lichtarge,O., Lieu,C., Liu,C., Liu,W., Louisged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapna,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Maxwell,E., Mcleod,M.P., Meador,M., Mel,G., Metzker,M., Miner,G., Minier,Z., Mitchell,T., Monabhat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokkwo,S., Oguh,M., Okunodu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pul,L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojoudokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoohtari,N., Stilson,I., Sodergren,E., Sonakke,T., Sparks,A., Stanley,A., Stone,H., Sutton,A., Svatek,A., Tabac,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,W., Wallington,S., Williams,G., Williamson,A., Wlezyg,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G., and Gibbs,R.
TITLE	Unpublished
JOURNAL	2 (bases 1 to 186656)
REFERENCE	Worley,K.C.
AUTHORS	Direct Submission
TITLE	Submitted (16-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
JOURNAL	On Dec 20, 2001 this sequence version replaced gi:15625866.
COMMENT	Genome Center Center: Baylor College of Medicine Center code: BCM Web site: http://www.hgsc.bcm.tmc.edu/ Contact: hgsc-help@bcm.tmc.edu Project Information Center project name: GEHL Center clone name: CH230-14D17 Summary Statistics Assembly program: Phrap, version 0.990329First call to findThrapList Consensus quality: 152607 bases at least Q40 Consensus quality: 159995 bases at least Q30 Consensus quality: 167102 bases at least Q20 Estimated insert size: 161475, sum-of-contigs estimation Quality coverage: 0x in Q20 bases; agarose-ef estimation Quality coverage: 2.2x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see <http://www.hgsc.bcm.tmc.edu/docs/genbankdraft.data.html>).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 61 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 6362: contig of 6362 bp in length
* 6363
* 6463: gap of unknown length
* 12730: contig of 6268 bp in length
* 12731
* 12831: gap of unknown length
* 12831
* 19424: contig of 6594 bp in length
* 19425
* 19525: gap of unknown length
* 26311: contig of 6787 bp in length
* 26312
* 26411: gap of unknown length
* 31962
* 31961: contig of 5550 bp in length
* 32061: gap of unknown length
* 32062
* 38719: contig of 6658 bp in length
* 38720
* 38819: gap of unknown length
* 38820
* 45557: contig of 6738 bp in length
* 45558
* 45657: gap of unknown length
* 45658
* 51953: contig of 6296 bp in length
* 51954
* 52053: gap of unknown length
* 52054
* 59692: contig of 7639 bp in length
* 59693
* 59793: gap of unknown length
* 59793
* 65470: contig of 5677 bp in length
* 65470
* 65569: gap of unknown length
* 65570
* 71068: contig of 5499 bp in length
* 71069
* 71169: gap of unknown length
* 75277: contig of 4109 bp in length
* 75278
* 75377: gap of unknown length
* 75378
* 80831: contig of 5454 bp in length
* 80832
* 80931: gap of unknown length
* 80932
* 84486: contig of 3555 bp in length
* 84487
* 84586: gap of unknown length
* 89240: contig of 4654 bp in length
* 89241
* 89340: gap of unknown length
* 89341
* 94016: contig of 4676 bp in length
* 94017
* 94116: gap of unknown length
* 94117
* 97616: contig of 3500 bp in length
* 97617
* 97716: gap of unknown length
* 97717
* 102256: contig of 4540 bp in length
* 102257
* 102356: gap of unknown length
* 102357
* 106697: contig of 4341 bp in length
* 106698
* 106797: gap of unknown length
* 106798
* 110404: contig of 3607 bp in length
* 110405
* 110504: gap of unknown length
* 110505
* 113150: contig of 2646 bp in length
* 113151
* 113250: gap of unknown length
* 113251
* 116916: contig of 3666 bp in length
* 116917
* 117016: gap of unknown length
* 117017
* 11837: contig of 2821 bp in length
* 11838
* 11937: gap of unknown length
* 11938
* 11937: gap of unknown length
* 11938
* 122653: contig of 2716 bp in length
* 122654
* 122753: gap of unknown length
* 122754
* 124554: contig of 1801 bp in length
* 124555
* 124654: gap of unknown length
* 124655
* 126969: contig of 2315 bp in length
* 126970
* 127069: gap of unknown length
* 127070
* 129204: contig of 2135 bp in length
* 129205
* 129304: gap of unknown length
* 129305
* 131556: contig of 2252 bp in length
* 131557
* 131656: gap of unknown length
* 131657
* 134102: contig of 2446 bp in length
* 134103
* 134202: gap of unknown length
* 134203
* 135651: contig of 1449 bp in length
* 135652
* 135751: gap of unknown length
* 135752
* 138622: contig of 2871 bp in length
* 138623
* 138722: gap of unknown length
* 140412: contig of 1690 bp in length

* 140413 140512: gap of unknown length
* 140513 142566: contig of 2054 bp in length
* 142567 142666: gap of unknown length
* 142667 143849: contig of 1183 bp in length
* 143850 143949: gap of unknown length
* 143950 146553: contig of 2704 bp in length
* 146554 146753: gap of unknown length
* 146754 147981: contig of 1228 bp in length
* 147982 148081: gap of unknown length
* 148082 150431: contig of 2350 bp in length
* 150432 150531: gap of unknown length
* 150532 152342: contig of 1811 bp in length
* 152343 152442: gap of unknown length
* 152443 153754: contig of 1312 bp in length
* 153755 153854: gap of unknown length
* 153855 155201: contig of 1347 bp in length
* 155202 155301: gap of unknown length
* 155302 157100: contig of 1799 bp in length
* 157101 157200: gap of unknown length
* 157201 159530: contig of 2330 bp in length
* 159531 159630: gap of unknown length
* 159631 161555: contig of 1925 bp in length
* 161556 161653: gap of unknown length
* 161656 163330: contig of 1675 bp in length
* 163331 163430: gap of unknown length
* 163431 164437: contig of 1007 bp in length
* 164438 164537: gap of unknown length
* 164538 166049: contig of 1512 bp in length
* 166050 167513: gap of unknown length
* 167514 167613: gap of unknown length
* 167614 168629: contig of 1016 bp in length
* 168630 168729: gap of unknown length
* 168730 169986: contig of 1257 bp in length
* 169987 170086: gap of unknown length
* 170087 171336: contig of 1350 bp in length
* 171337 171536: gap of unknown length
* 171537 173009: contig of 1473 bp in length
* 173010 173109: gap of unknown length
* 173110 174319: contig of 1210 bp in length
* 174320 174419: gap of unknown length
* 174420 175514: contig of 1095 bp in length
* 175515 17614: gap of unknown length
* 17615 17699: contig of 1185 bp in length
* 17699 176899: gap of unknown length
* 176800 177902: contig of 1003 bp in length
* 177903 178002: gap of unknown length
* 178003 179450: contig of 1448 bp in length
* 179451 179550: gap of unknown length

Query Match 26.6%; Score 315.6; DB 2; Length 186656;
Best Local Similarity 69.1%; Pred. No. 1.2e-64;
Matches 640; Conservative 0; Mismatches 226; Indels 60; Gaps 13:

QY 136 TCCGCCCGCCTCTCTGAGATTGCCCGAAGACACGCTGTGCTGCGCAGACG 195
|| ||||| || ||||| || || ||||| || || ||||| || || ||||| ||
DB 151939 TCTGCGCATGCTTCCTCGAAGCGTCCCAACACTGGGTATTTGCTGCGAGCG 151880
|| ||||| || ||||| || || ||||| || || ||||| || || ||||| ||
QY 196 GAGGCTGGGTGGAGAGGATGGCGGATTCACCGGATCGGAGCTGGTGAAC 255
|| ||||| || ||||| || || ||||| || || ||||| || || ||||| ||
DB 151879 CAGACCTGGGTGATGATGGAATTTGGCTGATTCGAACGGATATCTTCTTATTTGACT 151820
|| ||||| || ||||| || || ||||| || || ||||| || || ||||| ||
QY 256 ATCTCATCCCTGTTAGACCGATCGATATGTGC-AGAGTCAAGAAATTTGAT 314
|| ||||| || ||||| || || ||||| || || ||||| || || ||||| ||
DB 151819 GCCC-----ATAGACCAACCTGCTGTATGCAAGATTTCAAGAAATCTCCAT 151771
|| ||||| || ||||| || || ||||| || || ||||| || || ||||| ||
QY 315 CAGCTGCTGAGCAGTGGCTGTGACCTGTGACATGTAATCTGCAAAATGATGAGT 374
|| ||||| || ||||| || || ||||| || || ||||| || || ||||| ||
DB 151770 GAACTGCTGAGCAGTATCTGTAA--CTTGACATATTTACTTGAAATGATGAGT 151713
|| ||||| || ||||| || || ||||| || || ||||| || || ||||| ||
QY 375 CTTTACGCGGCATCGAGACCTTACAGGACGATGAGTGGAGGAGCCCGAGTA 434
|| ||||| || ||||| || || ||||| || || ||||| || || ||||| ||
DB 151712 CCTTGCTTTGTTATTTAGGATCCGTACAAACGATTAATGATAGAGATGCTGAGTA 151653

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 14:14:55 ; Search time 208.41 seconds
(without alignments)
9786.931 Million cell updates/sec

Title: US-09-898-216-2
Perfect score: 1188
Sequence: 1 GGCTTCTGGGACNACCGCT.....GGAGCAGATTTCCTGATT 1188

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1183	99.6	1188	19	AAV28867
2	1166	98.1	1322	20	AAV28867
3	1160	97.6	1398	22	ABA09225
4	1160	97.6	1398	22	AAK52683
5	1144.4	96.3	1337	20	AAV28867
6	1144	95.2	1416	22	AAK51699
7	1130.4	95.2	1244	22	AAH13961
8	1104.8	93.0	9098	22	AAV28867
9	516.6	43.5	691	23	AAV28867

10	516.4	43.5	567	22	AAK33192	Human CDNA clone r
11	516.4	43.5	567	22	AAH05115	Human CDNA clone (
12	515	43.4	518	22	AAI29477	Colon tumour relat
13	478.4	40.3	761	22	AAI24446	Human breast cance
14	465.6	39.2	566	22	AAK33925	Human CDNA clone r
15	406.4	34.2	1153	23	ABL05449	Drosophila melanog
16	392.8	33.1	2064	23	AAK94818	Human full-length
17	368.2	31.0	740	23	AAV57920	CDNA #596 encoding
18	327.8	27.6	421	22	AAI89469	Human polynucleoti
19	320.2	27.0	565	22	AAK92831	Human CDNA 3'-end
20	320.2	27.0	565	22	AAK94037	Human CDNA clone r
21	303.6	25.6	313	22	AAI23495	Human breast cance
22	299.4	25.2	301	22	AAV56972	Human breast cance
23	299.4	25.2	301	22	AAI17542	Human breast cance
24	280.8	23.6	1677	21	AAAI2730	DNA encoding of a
25	280.6	23.6	1507	21	AAK49579	Arabidopsis thalia
26	269.4	22.7	1507	21	AAK50093	Arabidopsis thalia
27	269.4	22.7	1519	21	AAK38968	Arabidopsis thalia
28	269.4	22.7	1548	21	AAK43010	Arabidopsis thalia
29	262.6	22.1	1387	21	AAK40332	Arabidopsis thalia
30	259.4	21.8	417	22	AAI15602	Human breast cance
31	256.8	21.6	1682	22	AAV5141	CDNA encoding nove
32	179.4	15.1	253	15	AAQ76727	Human genome fragm
33	173.8	14.6	190	21	AAI17814	Human gene express
34	171.2	14.4	747	20	AAI215045	Human gene express
35	159.4	13.4	202	21	AAI24655	Human 5' EST Isola
36	159.2	13.4	207	23	AAV58068	CDNA #744 encoding
37	158.2	13.3	3369	23	ABL18422	Drosophila melanog
38	158.2	13.3	5058	23	ABL05448	Drosophila melanog
39	156.8	13.2	607	22	AAI14633	Human breast cance
40	156.8	13.2	948	21	AAI21550	N. meningitidis pa
41	156.8	13.2	948	21	AAI21553	N. meningitidis pa
42	156.8	13.2	948	21	AAI1265	N. meningitidis pa
43	156.8	13.2	948	21	AAI1265	N. meningitidis pa
44	156.8	13.2	948	21	AAI23751	N. meningitidis pa
45	156.8	13.2	948	21	AAI23751	Neisseria meningit

ALIGNMENTS

RESULT 1	AAV28867	standard; DNA: 1188 BP.
XX	AAV28867;	
XX	03-AUG-1998	(first entry)
XX	Human integral membrane protein encoding DNA.	
XX	Human: integral membrane protein; IMP; cancer; anaemia; prostate;	
XX	breast; pancreatic; tumour; ion transport; ss.	
XX	Homo sapiens.	
XX	OS	
XX	Key	Location/Qualifiers
XX	CDS	64..1134
XX	FT	/*tag=
XX	FT	/product= "Integral membrane protein"
XX	FT	/transl_except= (pos:490..492,aa:Xaa)
XX	FT	/transl_except= (pos:1099..1101,aa:Xaa)
XX	FT	/note= "Xaa= unknown"
XX	US5763589-A.	
XX	09-JUN-1998.	
XX	09-JAN-1997;	97US-0781562.
XX	09-JAN-1997;	97US-0781562.
XX	(INCY-) INCYTE PHARM INC.	

XX Goli SK, Hillman JL;
 XX MPI: 1998-347418/30.
 DR P-PSDB: AAM57232.
 XX
 PT DNA encoding human Integral membrane protein - useful for producing
 PT recombinant protein, for treatment of anaemia and cancer
 XX
 PS Claim 3: Column 37-40; 33pp; English.
 XX
 CC The present sequence encodes human Integral membrane protein (IMP).
 CC IMP may be administered to a subject to treat disorders associated
 CC with abnormal ion transport or membrane conductance as well as a
 CC variety of tumours, e.g. haemolytic anaemias and prostate, breast and
 CC pancreatic tumours. A vector capable of expressing IMP, or a fragment
 CC or a derivative thereof, may also be administered to a subject to treat
 CC the haemolytic anaemias and prostate, breast and pancreatic tumours.
 XX
 SQ Sequence 1188 BP: 268 A; 316 C; 362 G; 237 T; 5 other:

Query Match 99.6%; Score 1183; DB 19; Length 1188;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTTGTGGAGNACGCTCCGCTCGTCTGTTGTTCCGAGAGTCGCTCGCGGTGG 60
 Db 1 ggcctctggagacacgcctccgctcgtctgcttgctcggaggtcgtcgcggtg 60
 QY 61 GAATGCTGGCG 120
 Db 61 gaaatgctggcg 120
 QY 121 GGCTTGTGGCG 180
 Db 121 ggcctctggcg 180
 QY 181 TTCTGTGCGCGAGAGAGCGCTGGGTGGTGGAGGAAATGGCGCATTCACCGATCCTG 240
 Db 181 ttctgtgctgc 240
 QY 241 GAGCCTGGTTGACATTCCTCTATCCCTGTGTAGACGCGATCCGATATGTGCAGACTTC 300
 Db 241 gagcctgggttgacatctccttctgtgttagacgcgcgcgcgcgcgcgcgcgcgc 300
 QY 301 AAGGAAATTTGATCAACAGTCTGAGAGAGTGGGTGATCTCGACATTAAGTAACTGTG 360
 Db 301 aaggaaatttgatcaaacagtctgagagagtggtgatctcgacatttaagttaactgt 360
 QY 361 CAATTCATGAGTCTTTTACCTCGCATGTGACCCCTTACAGGCAAGCTACGGTGTG 420
 Db 361 caatctcatgagtccttttacctctgcatgtgaccccttacaaggcaagctaacgtgtg 420
 QY 421 GAGAGACCTGAGTATGCGCTTACCCAGCTACTCAAAACATGAGATCAGAGCTCGGC 480
 Db 421 gagagacctgagtatgctgttacccttaccctgcatcaaaacatgagatcagagctcgc 480
 QY 481 AAATCTCTNTGACAAAGTCTTCCGCGAAGAGTCCCTGAATGCGAGATTTGTGAT 540
 Db 481 aaatctctntgacaaagtcttcccggaagagtccttgaatgctgagatctgtgat 540
 QY 541 GCCATCAACAGTCTGCTGCTGGGTATCCGCTGCCCTNCGTTATGATCAAGAGAT 600
 Db 541 gccatcaaacagtctgctgctgggtatccgctgccctncttattgatcaagagat 600
 QY 601 ATTCATGCGCACCGCGGTGAAGAGTCTATGCAANTGCAAGTGGAGCGAGCGGGG 660
 Db 601 attcattgctgcacccggtgaagagcttatgcaantgcaagtggagcgagcgagg 660
 QY 661 AAGGCGGCACAGTTTGAAGTCTGAGGAGACCGAGAGTGGCGCATATGTGGACAGAA 720
 Db 661 aaggcggcacagtttgaagtctgagagacccgagagtgctgcacatattgtgcagaa 720

QY 721 GGGAAGAAACAGGCCAGATCTGGGCTTCGAGACAGAAAAGCTGAACAGATAATCAG 780
 Db 721 gggaagaagacagggccagatctgggcttcgagacagaaaagctgacagataaatcag 780
 QY 781 GCAGCAGAGAGGCGCAGTTCGTCGCGCAAGGCGCAAGGCTGAAGCTATTTGCA 840
 Db 781 gcagcagagagggcagtgcttctgagagggcgaagctgaagctaatctatctga 840
 QY 841 ATCTGTGCTGAGCTCTACACATTAATGAGATGCGAGCGTCTGACTGTGGCC 900
 Db 841 atctgtctgagctctacacattaatgagatgctgagcgcttctgactgtggcc 900
 QY 901 GAGCAGTATTCAGCGCGCTTCTCCAACTGGCCCAAGACTCAACACTATCTACTGCC 960
 Db 901 gagcagtatttcagcgcgcttctccaaactggcccaagctcaaacactatctactgcc 960
 QY 961 TCCAACCTGGCGCATGTACACAGCATGTGCTCAGGCCATGGGTGTTATGAGCCCTC 1020
 Db 961 tccaacctggcgcatgtacacagcatgtgctcagggccatgggtgttatgagccctc 1020
 QY 1021 ACCAAGCGCCGAGTGCAGGAGCTCCAGACTCAGTCTCAGTGGAGCAGAGATGTC 1080
 Db 1021 accaagcgccgagtgctcagagctccagactcagctcagtgaggagcagagatgctc 1080
 QY 1081 CAGGATACAGATGCAAGTNTTATGAGAACTTATGATCAGATGATGATGAGC 1140
 Db 1081 caggatcacagatgcagagntttatgagaaacttattgatcagatgattgagtcagtc 1140
 QY 1141 TGGGCTTGGCAGGAGCTGCGGAGCAAGCAAGCAAGATTTTCTGATTT 1188
 Db 1141 tgggcttggcagagctgctggagcaagcaagcaagattttctgatttt 1188

RESULT 2
 AAX04402
 ID AAX04402 standard; DNA; 1322 BP.
 XX
 AC AAX04402;
 XX
 DT 13-Apr-1999 (first entry)
 XX
 DE Human secreted protein gene 35 clone HTXCS21.
 XX
 KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 OS Homo sapiens.
 XX
 PN W03856804-A1.
 XX
 PD 17-DEC-1998.
 XX
 PF 11-JUN-1998; 98MO-US12125.
 XX
 PR 02-OCT-1997; 97US-0061060.
 PR 13-JUN-1997; 97US-0049547.
 PR 13-JUN-1997; 97US-0049548.
 PR 13-JUN-1997; 97US-0049549.
 PR 13-JUN-1997; 97US-0049550.
 PR 13-JUN-1997; 97US-0049550.
 PR 13-JUN-1997; 97US-0049607.
 PR 13-JUN-1997; 97US-0049608.
 PR 13-JUN-1997; 97US-0049609.
 PR 13-JUN-1997; 97US-0049610.
 PR 13-JUN-1997; 97US-0049611.
 PR 13-JUN-1997; 97US-0050566.

KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; actlyin;
KW inhibin; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asplenia; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antinflammatory; antiarthritic; haemostatic; antiatherosclerotic;
KW cystostatic; osteopathic; vasotropic; cardiant; virocidic; antibacterial;
KW antifungal; vulnery; antitumor; ss.
XX Homo sapiens.
XX Mo200157188-A2.
XX 09-AUG-2001.
XX 05-FEB-2001; 2001MO-US03800.
XX 03-FEB-2000; 2000US-0496914.
XX 27-APR-2000; 2000US-0560875.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT;
XX WPI: 2001-457740/49.
XX P-PSDB; AB111981.
XX Human proteins and DNA encoding sequences useful for preventing,
XX treating or ameliorating a medical condition in a mammalian subject
XX e.g. arthritis and cancer -
XX Claim 1: Page 851; 1963pp; English.
XX
XX Sequences AB10981-AB12330 represent 1350 novel human polypeptides, and
XX sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
XX invention also relates to vectors and recombinant host cells comprising a
XX nucleotide of the invention, methods of producing the novel polypeptides,
XX antibodies against the polypeptides, methods of detecting the nucleotides,
XX or polypeptides in a sample, and methods of identifying compounds which
XX bind to polypeptides of the invention. Although novel, many of the
XX polypeptides of the invention have homology to known proteins, thereby
XX giving an insight into their probable biological activities, and hence
XX potential therapeutic applications. The polypeptides of the invention may
XX have various activities, including cytokine, cell proliferation or cell
XX differentiation activities; stem cell growth factor activity;
XX haematopoiesis regulatory activity; tissue growth activity;
XX immunomodulatory activity; activin- or inhibin-related activities;
XX chemotactic or chemokinetic activities; haemostatic, thrombotic or
XX thrombolytic activities; receptor or ligand activities; or may be
XX involved in oncogenesis, cancer cell proliferation or metastasis.
XX Depending on their biological activities, polypeptides and nucleotides of
XX the invention are useful for preventing, treating or ameliorating medical
XX conditions, e.g., by protein or gene therapy. Such conditions include
XX cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
XX disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
XX proliferative retinopathy, atherosclerosis, coronary heart disease,
XX arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
XX vascular growth. Polypeptides involved with tissue regeneration and
XX repair (or nucleic acids encoding them) may be used to promote wound
XX healing (e.g., of burns, incisions and ulcers), while those with
XX immunomodulatory activities may be used in the treatment of viral,
XX bacterial and fungal infections in addition to immune disorders.
XX Polypeptides with growth factor activity may be used in cell cultures to
XX promote cell growth. For example, such polypeptides may be used to
XX manipulate stem cells in culture to give rise to neuroepithelial cells
XX that can be used to augment or replace cells damaged by illness,
XX autoimmune disease or accidental damage. The polypeptides and nucleotides
XX may also be used in the diagnosis of the above conditions, and in drug

CC screening techniques. The present sequence represents a cDNA encoding a
CC novel human polypeptide of the invention.
XX
SQ Sequence 1398 BP; 307 A; 408 C; 359 G; 324 T; 0 other;
Query Match 97.6%; Score 1160; DB 22; Length 1398;
Best Local Similarity 99.4%; Fred. No. 0;
Matches 1182; Conservative 0; Mismatches 5; Indels 2; Gaps 2;
1 GGCCTCTGGAGACNNACCGCTCCGTCGTCCTGTGTTCCGAGGTCCGCGCGTGG 60
|||||
DB 1393 GCGTCTGGAGACGACGCGCTCCGCTCCTGTTGTTCCGAGGTCCGCGCGTGG 1334
61 GAATGCTGGCG 120
|||||
DB 1333 GAATGCTGGCG 1275
121 GCGTCTGGCG 179
|||||
DB 1274 GCGTCTGGCG 1215
180 GTTCG 239
|||||
DB 1214 GTTCG 1155
240 GGAGCGCTGTTGAACATCCTCATCCCTGTGTAGACCGGATCCGATATGTCAGAGTCT 299
|||||
DB 1154 GGAGCGCTGTTGAACATCCTCATCCCTGTGTAGACCGGATCCGATATGTCAGAGTCT 1095
300 CAAGGAATTTGTCATCAACGTCGCTAGCAGTCGCTGTGACTCTCGAACAATGTAATCT 359
|||||
DB 1094 CAAGGAATTTGTCATCAACGTCGCTAGCAGTCGCTGTGACTCTCGAACAATGTAATCT 1035
360 GCAATCGATGAGTGCCTTACCTGGCGCATCATGACCCCTTACAAGCAAGCTACGCTCT 419
|||||
DB 1034 GCAATCGATGAGTGCCTTACCTGGCGCATCATGACCCCTTACAAGCAAGCTACGCTCT 975
420 GGAGGACCTGTGATGCGGTGACCCGAGTACGTAACACCATGATGATCAAGCTCGG 479
|||||
DB 974 GGAGGACCTGTGATGCGGTGACCCGAGTACGTAACACCATGATGATCAAGCTCGG 915
480 CAATCTCTGTGACCAAGTCTTCCGGGAGCGGAGTCCCTGAAATGCAAGCATTTGTGA 539
|||||
DB 914 CAATCTCTGTGACCAAGTCTTCCGGGAGCGGAGTCCCTGAAATGCAAGCATTTGTGA 855
540 TGGCATCAACCAAGTCTGTCAGTCTGGGGTATCCCTGCTTCGTTATGAGATCAAGA 599
|||||
DB 854 TGGCATCAACCAAGTCTGTCAGTCTGGGGTATCCCTGCTTCGTTATGAGATCAAGA 795
600 TATCATGTGCCACCCCGGGTGAAGAAGTCTATGACATGACAGTGGAGCGAGCGCG 659
|||||
DB 794 TATCATGTGCCACCCCGGGTGAAGAAGTCTATGACATGACAGTGGAGCGAGCGCG 735
660 GAACCGGCCACAGTCTGAGAGTCAAGGAGACCGAGAGTGGCGCATGAATGTGCACA 719
|||||
DB 734 GAACCGGCCACAGTCTGAGAGTCAAGGAGACCGAGAGTGGCGCATGAATGTGCACA 675
720 AGGAGAAACAGAGGCGCAATCCTGGGCTCCGAAAGCAAGAAAGGCTGAACAGATTAATCA 779
|||||
DB 674 AGGAGAAACAGAGGCGCAATCCTGGGCTCCGAAAGCAAGAAAGGCTGAACAGATTAATCA 615
780 GCGCAGCAGAGAGCGCAGTGCAGTCTTGGCGAAGGCCAAGGCTTAAGCTGAAGCTTTGG 839
|||||
DB 614 GCGCAGCAGAGAGCGCAGTGCAGTCTTGGCGAAGGCCAAGGCTTAAGCTGAAGCTTTGG 555
840 AATCTGGCTGCGAGTCTGACACACATATGAGATGAGAGCGTTCACGCTGTGGC 899
|||||
DB 554 AATCTGGCTGCGAGTCTGACACACATATGAGATGAGAGCGTTCACGCTGTGGC 495
900 CGAGCAGTATGTAGAGCGCTTCCAAAGCTGCGCAAGGACTCCACACTATCTACTGCG 959
|||||
DB 494 CGAGCAGTATGTAGAGCGCTTCCAAAGCTGCGCAAGGACTCCACACTATCTACTGCG 435

OY 960 CTCACACCCGTGGCATGTCCAGCATGTTGGTCCAGGCGCATGGTGTATATGGAGCCCT 1019
|||||
DB 434 CTCACACCCCTGGGGATGTCCACGACATGTGGCTCAAGCCATGGTATATGGAGCCCT 375
|||||
OY 1020 CACCAAGGCCCGAGTGGCAGGAGCTCCAGACTCTCTCCAGTGGAGCAGCAGATGT 1079
|||||
DB 374 CACCAAGGCCCGAGTGGCAGGAGCTCCAGACTCTCTCCAGTGGAGCAGCAGATGT 315
|||||
OY 1080 CCAGGCTACGATGCAAGTATTGATGAGAACCTTGATCGATCAAGTGAAGTGGAG 1139
|||||
DB 314 CCAGGCTACGATGCAAGTATTGATGAGAACCTTGATCGATCAAGTGAAGTGGAG 255
|||||
OY 1140 CTGGGCTTGGCCAGGAGCTGGGGACAGGAAGCAGATTTTCTGATT 1188
|||||
DB 254 CTGGGCTTGGCCAGGAGCTGGGGACAGGAAGCAGATTTTCTGATT 206
|||||
RESULT 4
AAK52683/c
ID AAK52683 standard; cDNA: 1398 BP.
XX AAK52683;
AC
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 2212.
XX
XX Human: cytokine; cell proliferation; cell differentiation; gene therapy;
KM vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KM tissue growth factor; immunomodulatory; cancer; leukaemia;
KM nervous system disorder; arthritis; inflammation; ss.
OS Homo sapiens.
XX
XX WO200157190-A2.
PN
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
XX
PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX WPI: 2001-476283/51.
DR P-PSDB; AAM79550.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
XX
PS Claim 1; Page 4572; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAW80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, Leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
XX

Sequence 1398 BP; 307 A; 408 C; 359 G; 324 T; 0 other;

Query Match 97.6%; Score 1160; DB 22; Length 1398;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1182; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

OY 1 GGCCTTGGGAGCAGACCCGCCCTGCTCTCTGCTTCCGGAGAGTGGCGCGGCTGG 60
|||||
DB 1393 GGCCTTGGGAGCAGACCCGCCCTGCTCTCTGCTTCCGGAGAGTGGCGCGGCTGG 1334
|||||
OY 61 GAATGCTGGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
|||||
DB 1333 GAATGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1275
|||||
OY 121 GGCCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 179
|||||
DB 1274 GGCCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1215
|||||
OY 180 GTTCGTCGCGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 239
|||||
DB 1214 GTTCGTCGCGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1155
|||||
OY 240 GAGAGCCGTGTTGAAACATCTCATCCCTGTGTTAGACCGGATCCGATATGTGAGAGTCT 299
|||||
DB 1154 GAGAGCCGTGTTGAAACATCTCATCCCTGTGTTAGACCGGATCCGATATGTGAGAGTCT 1095
|||||
OY 300 CAAGGAATTTGTCATCAAGTGGCTGAGCAGTGGCTGTGACTGTGACAAATGTAATCT 359
|||||
DB 1094 CAAGGAATTTGTCATCAAGTGGCTGAGCAGTGGCTGTGACTGTGACAAATGTAATCT 1035
|||||
OY 360 GCAATGATGAGAGTCCCTTACCTGCGCATCATGAGACCCCTTACAGGCAAGCTACGCTGT 419
|||||
DB 1034 GCAATGATGAGAGTCCCTTACCTGCGCATCATGAGACCCCTTACAGGCAAGCTACGCTGT 975
|||||
OY 420 GAGGAGCCCTGAGTATGCGCTCAACCCAGTACGTCMAAACCAACATGAGATCAGAGCTCG 479
|||||
DB 974 GAGGAGCCCTGAGTATGCGCTCAACCCAGTACGTCMAAACCAACATGAGATCAGAGCTCG 915
|||||
OY 480 CAAACTCTCTNTGACAAAGTCTTCCGGGAACGGAGTCCCTGATATCCAGCATTTGTGGA 539
|||||
DB 914 CAAACTCTCTNTGACAAAGTCTTCCGGGAACGGAGTCCCTGATATCCAGCATTTGTGGA 855
|||||
OY 540 TGGCATCAACCAAGCTGCTGAGTGGGTATCCGCTGGCTGCTGCTGCTGCTGCTGCTGCT 599
|||||
DB 854 TGGCATCAACCAAGCTGCTGAGTGGGTATCCGCTGGCTGCTGCTGCTGCTGCTGCTGCT 795
|||||
OY 600 TATTCATGTCCACCCCGGGGTGAAGAGTCTATGCAATGATGAGTGGAGCAGAGCGGCG 659
|||||
DB 794 TATTCATGTCCACCCCGGGGTGAAGAGTCTATGCAATGATGAGTGGAGCAGAGCGGCG 735
|||||
OY 660 GAAACGGGCGACAGTTCTAGAGTGAAGGAGCCGAGAGTGGGCCATCAATGTGCGAGA 719
|||||
DB 734 GAAACGGGCGACAGTTCTAGAGTGAAGGAGCCGAGAGTGGGCCATCAATGTGCGAGA 675
|||||
OY 720 AGGGAAGAAACAGGCGGCGGAGTCCGCGGCTCGGAAGCAAGAAAGCGTAACAGATAATTC 779
|||||
DB 674 AGGGAAGAAACAGGCGGCGGAGTCCGCGGCTCGGAAGCAAGAAAGCGTAACAGATAATTC 615
|||||
OY 780 GGCAGCAGAGAGGCGGCGGAGTCTGCGGAGGCGCAAGGCTTAAAGCTAAAGCTATTTCG 839
|||||
DB 614 GGCAGCAGAGAGGCGGCGGAGTCTGCGGAGGCGCAAGGCTTAAAGCTAAAGCTATTTCG 555
|||||
OY 840 AATCTGCTGACAGCTCTGACACAAATATGAGATGACAGAGCTTCACTGACTGTGGC 899
|||||
DB 554 AATCTGCTGACAGCTCTGACACAAATATGAGATGACAGAGCTTCACTGACTGTGGC 495
|||||

Oy	900	CGACGATGATGTGACGCCGCTTTCCTCCAAACCTGGCCAGACACTCCACACTATCTCTACTGCC	959
Db	494	CGACGATGATGTGACGCCGCTTTCCTCCAAACCTGGCCAGACACTCCACACTATCTCTACTGCC	435
Oy	960	CTCCAAACCTTGGCCATGTTCACACGATGATGGCTCAGGCCATGGGTATATGAGACCCCT	1019
Db	434	CTCCAAACCTTGGCCATGTTCACACGATGATGGCTCAGGCCATGGGTATATGAGACCCCT	375
Oy	1020	CACCAAAAGCCCACTGCCAGGAGACTCCAGACTCTCTCAGTGGGAGCAGCAGAGATGT	1079
Db	374	CACCAAAAGCCCACTGCCAGGAGACTCCAGACTCTCTCAGTGGGAGCAGCAGAGATGT	315
Oy	1080	CCAGGCTACAGATCATGATNTTGATGAGAACTTGATCGATCAGATAGTTAGTGGAG	1139
Db	314	CCAGGCTACAGATCATGATNTTGATGAGAACTTGATCGATCAGATAGTTAGTGGAG	255
Oy	1140	CTGGCCTTGGCCAGGAGTCTGGGGACAGAAAGCAGATTTTCCTGATT	1188
Db	254	CTGGCCTTGGCCAGGAGTCTGGGGACAGAAAGCAGATTTTCCTGATT	206
RESULT 5			
AAK0345	AAK0345 standard; DNA: 1337 BP.		
AC	AAK0345:		
XX			
DT	13-APR-1999 (first entry)		
XX			
DE	Human secreted protein gene 35 clone HTXCS21.		
XX			
KW	Human; secreted protein; fusion protein; gene therapy; protein therapy;		
KW	diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;		
KW	developmental abnormality; foetal deficiency; blood; allergy; renal; ds;		
KW	immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;		
KW	inflammation; ischemic shock; Alzheimer's disease; restenosis; AIDS;		
KW	cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus		
KW	osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;		
KW	endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.		
XX			
OS	Homo sapiens.		
XX			
PN	W09856804-A1.		
PD	17-DEC-1998.		
XX			
PF	11-JUN-1998; 98WO-US12125.		
XX			
PR	02-OCT-1997;	97US-0061060.	
PR	13-JUN-1997;	97US-0049547.	
PR	13-JUN-1997;	97US-0049548.	
PR	13-JUN-1997;	97US-0049549.	
PR	13-JUN-1997;	97US-0049550.	
PR	13-JUN-1997;	97US-0049606.	
PR	13-JUN-1997;	97US-0049607.	
PR	13-JUN-1997;	97US-0049608.	
PR	13-JUN-1997;	97US-0049609.	
PR	13-JUN-1997;	97US-0049610.	
PR	13-JUN-1997;	97US-0049611.	
PR	13-JUN-1997;	97US-0050566.	
PR	13-JUN-1997;	97US-0050901.	
PR	13-JUN-1997;	97US-0052989.	
PR	08-JUL-1997;	97US-0051919.	
PR	18-AUG-1997;	97US-0055984.	
PR	12-SEP-1997;	97US-0058665.	
PR	12-SEP-1997;	97US-0058668.	
PR	12-SEP-1997;	97US-0058669.	
PR	12-SEP-1997;	97US-0058750.	
PR	12-SEP-1997;	97US-0058971.	
PR	12-SEP-1997;	97US-0058972.	
PR	12-SEP-1997;	97US-0058975.	
PR	02-OCT-1997;	97US-0060834.	
PR	02-OCT-1997;	97US-0060841.	

[illegible]

[illegible]

RESULT	6
AAK51699	
ID	AAK51699 standard; cDNA; 1416 BP.
XX	
AC	AAK51699;
XX	
DT	06-NOV-2001 (first entry)
XX	
DE	Human polynucleotide SEQ ID NO 244.
XX	
KW	Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW	tissue growth factor; immunomodulatory; cancer; leukaemia;
KW	nervous system disorder; arthritis; inflammation; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200157190-A2.
XX	
PD	09-AUG-2001.
XX	
PF	05-FEB-2001; 2001WO-US04098.
XX	

03-FEB-2000; 2000US-0496914.
27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.

(HYSE-) HYSEQ INC.

XX
PA
XX
PI Tang YF, Liu C, Dermanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QH, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
DR WPI: 2001-476283/51.
DR P-PSDB: AAM78556.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
PS Claim 1; Page 1135-1136; 6221pp: English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating
CC cytokine, cell proliferation or cell differentiation on which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities
e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
XX Sequence 1416 BP; 355 A; 364 G; 410 G; 307 T; 0 other;

[illegible]

QY 389 TCATGACCCCTTACAAAGCAGCTACGCTGTGAGAGACCCCTGATGTCCTGACCCAGC 448
|||||
Db 360 tcattgacccttaacaaggaaagtaagctgtgtgagagacccttgatgctgcgcacaccagc 419
QY 449 TAGCTCAAAACATCATGAGATCAGAGCTCGCAAACTCTCTGTGCAAAATCTTCCGGG 508
|||||
Db 420 tagctcaaaacacacatagatcagagctcgcaaaactctctctggaacaaagctctccgg 479
QY 509 AACGGGAGTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 568
|||||
Db 480 aacgggagtccttgatgacagcatgtgtgagatgcaacaaagctctgactgtctgg 539
QY 569 GTATCCGCTGCTTCGCTTATGATATCAAGATATCATGCTGCTGCTGCTGCTGCTGCTG 628
|||||
Db 540 gatacgcgcctcctcgcttctgagatcaagaatatacctatgtgcaaccccggtgaaagat 599
QY 629 CTATGACATGACAGCTGAGGAGCAGACGGCGGAAACGGGCGCACAGTTCTAGAGTCTAGG 688
|||||
Db 600 ctatgcagatgcaggtgtgagggcagagcggcgaacggcgccacagctctcagagctcag 659
QY 689 GGACCCGAGAGTCGGCATCATGATGTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 748
|||||
Db 660 ggaacccgagatgcgcacatcaatgtgcagaaaggaaagaaacagagccagatctctggcct 719
QY 749 CCGAAGCAGAAAGGCTGAAACAGATTAATCAAGCAGACAGAGAGGAGGAGGAGGAGGAGG 808
|||||
Db 720 ccgaagcagaaagagctgagacagataaatacaagcagcagcagcagcagcagcagcagc 779
QY 809 CGAAGCCCAAGGCTAAAGCTGAAGCTATTGCAATCTGCTGCTGCTGCTGCTGCTGCTG 868
|||||
Db 780 cgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 839
QY 869 ATGGAGATGACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 928
|||||
Db 840 atggagatgacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 899
QY 929 TGGCAAGAGCTCAACACTATCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 988
|||||
Db 900 tggcgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 959
QY 989 TGGCTCAGCCCATGGGTATATGAGAGCCCTCAACCAAGCCCAAGTGCAGAGCTCCAG 1048
|||||
Db 960 tggctcagcccatgggtatattgagagccctcaaccccaagcagcagcagcagcagcagcag 1019
QY 1049 ACTGACTCTCCAGTGGGAGCAGAGATGTCAGGCTGACATGCAAGTNTTGTATGAGAG 1108
|||||
Db 1020 actgactctccagtgaggacagcagcagcagcagcagcagcagcagcagcagcagcagc 1079
QY 1109 AACTGATCAGATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1168
|||||
Db 1080 aactgactcagatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1139
QY 1169 GGAAGCAGATTTTCTGATT 1188
|||||
Db 1140 ggaagcagatttctcgtalt 1159

RESULT 8
AAS44953
ID AAS44953 standard: cDNA: 9098 BP.
XX
AC AAS44953:
XX
DT 18-DEC-2001 (first entry)
XX
DE cDNA encoding novel human secretory protein, Seq ID No 34.
XX
KM Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
KM Ischemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
KM Transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
KM Amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
KM Ulcer; osteoporosis; bone degenerative disorder; periodontal disease;

KW gut protection; lung; liver fibrosis; immune deficiency; infection;
KM severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
KM multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
KM fertility; analgesic; pain; antigen; ss.
OS Homo sapiens.
XX
PN WO200166689-A2.
XX
PD 13-SEP-2001.
XX
PF 05-MAR-2001; 2001WO-US04942.
XX
PR 07-MAR-2000; 2000US-0519705.
PR 19-MAY-2000; 2000US-0574454.
PR 17-JUN-2000; 2000US-0596193.
PR 14-JUL-2000; 2000US-0618847.
PR 19-SEP-2000; 2000US-0665363.
PR 20-OCT-2000; 2000US-0693267.
XX
PA (HSE-) HSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
PI Zhao QA, Yang Y, Dirmanc RT, Zhang J, Chen R, Xue AJ, Wang J;
XX
DR MPI: 2001-589934/66.
XX
PT P-PSDB: AAI28053.
PT Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis and treatment of
PT cancer, neurological, inflammatory, and autoimmune disorders -
XX
PS
Claim 1: SEQ ID No 34; 107pp; English.
XX
CC The invention relates to novel isolated human secreted polypeptides (I)
CC and polynucleotides (II). (I) and (II) are useful for treating
CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
CC ischemia-reperfusion injury, shock, sepsis, immune responses, and is
CC involved in increasing haematopoiesis, stem cell survival, bone growth
CC and remodeling. (I), (II) and modulators of (II) are useful for
CC prophylaxis or treatment of one or more cancers. (II) is also useful for
CC creating transgenic animals useful for studying the in vivo activities of
CC the polypeptide as well as for studying modulators of the polypeptides.
CC (I) induces the proliferation of neural cells and regeneration of nerve
CC and brain tissue and is useful for the treatment of central and
CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
CC activity, regulation of haematopoiesis and is useful for treating myeloid
CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
CC tissue growth and in tissue repair, healing of burns, incisions,
CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
CC disorders, or periodontal disease. Furthermore, (I) is also useful for
CC gut protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, various immune deficiencies and
CC disorders including severe combined immunodeficiency (SCID), bacterial or
CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
CC reactions and conditions, such as asthma or other respiratory problems.
CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
CC fertility, metabolism, catabolism, anabolism, storage or elimination of
CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
CC analgesic effects or other pain reducing effects, immunoglobulin like
CC activity and can act as an antigen in a vaccine composition to raise an
CC immune response. AAS44920-AAS45295 represent novel human secreted protein
CC coding sequences of the invention.
XX
SQ Sequence 9098 BP; 1918 A; 2538 C; 2590 G; 2052 T; 0 other;

Query Match 93.0%; Score 1104.8; DB 22; Length 9098;
Best Local Similarity 99.3%; Pred. No. 2, 8e-303;

Matches 1128; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

```

Oy 1 GACCTTGTGGAGACNACCGCTCCGCTGCTGCTGTTCCGAGGCTCGCTGCGGGGTGG 60
    |||
Db 12 ggctctctggagagacagctccgctcgtctcgtctcgtctcgtctcgtctcgtctcgt 71
Oy 61 GAAATGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
    |||
Db 72 gaatctgctgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcg 130
Oy 121 GCGCTTCTGGCCCCCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
    |||
Db 131 ggctctctgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcg 190
Oy 180 GTTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 239
    |||
Db 191 gctgctgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcg 250
Oy 240 GGAGCGCTGGTTGAACATCTCATCTCTGTGTAGACCGGATCCGATATGTGACAGTCT 299
    |||
Db 251 ggagcgctgcttgaaacatccatccatccatccatccatccatccatccatccatcc 310
Oy 300 CAGGAAATTTGTCATCAACAGTGGCTGAGCAGTGGCTGAGCTGAGCTGAGCTGAGCT 359
    |||
Db 311 caagaaatctgcatcaacacgctgctgagcagctgagctgagctgagctgagctgag 370
Oy 360 CCAATCGATGAGTCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 419
    |||
Db 371 gcaatcgatgagctgagctgagctgagctgagctgagctgagctgagctgagctgag 430
Oy 420 GGAGGACCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 479
    |||
Db 431 ggagcgctgcttgaaacatccatccatccatccatccatccatccatccatccatcc 490
Oy 480 CAAACTCTCTTNGCAAAAGTCTTCCGGGAGAGGAGTCCCTGATGCTGAGCTGATGTA 539
    |||
Db 491 caaactctctcggagaaagctctccggagacgagctccggagacgagctccggagac 550
Oy 540 TGCCATCAACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 599
    |||
Db 551 tgcctcaacacagctgctgagctgctgagctgctgagctgctgagctgctgagctgag 610
Oy 600 TATCCATGTGCCACCCCGCGGTGAAAGTCTATGCAAGATGAGCTGAGGAGAGCGGCG 659
    |||
Db 611 tatccatgtccaccccggtgaaagctctgcaagctgcaagctgcaagctgcaagctgca 670
Oy 660 GAAAGGGGCGACAGTTTAAAGTCTGAGGGGACCGGAGAGTGGCGGATTAATGTGGCA 719
    |||
Db 671 gaagcgccacagctctgagctgagctgagctgagctgagctgagctgagctgagctgag 730
Oy 720 AGGGAAGAAACAGCGCGCATCTGGCTCCGAGACAGAAAGCTGAGCAGATTAATCA 779
    |||
Db 731 agggagaaacagcgccagctctgagctgagctgagctgagctgagctgagctgagctgag 790
Oy 780 GGCAGCAGAGAGCGCCAGTGTGCTGCGGAGGCCAAGGCTAAAGCTGAAGCTATTGCG 839
    |||
Db 791 ggcaagcagagagcgagctgagctgagctgagctgagctgagctgagctgagctgag 850
Oy 840 AATCTGGCTGACGCTGACACATTAATGAGATGACAGAGCTGACCTACTCTGCG 899
    |||
Db 851 aatctgctgagctgagctgagctgagctgagctgagctgagctgagctgagctgagctgag 910
Oy 900 CGAGCAGTATGACAGCGCTTCTCCAAACTGCGCAAGAGCTCAAGACTATCTCTACTGCG 959
    |||
Db 911 cgagcagctatgctcagcgctctcctcaaacctgagcagcagcagcagcagcagcagcagc 970
Oy 960 CTGCAACCTTGGCGATGTACACGAGATGGTGGCTCAGGCGCATGGCTGATATGAGCCCT 1019
    |||
Db 971 ctccaacccctgagctgagctgagctgagctgagctgagctgagctgagctgagctgag 1030
Oy 1020 CACCAAAACCCCAAGTGCAGGAGTCTGACACTCTCTCAATGGGAGGAGCAGAGATGT 1079
    |||
Db 1031 caccaaaagcccaagctgagcagagctccagctcctcagctgagctgagctgagctgagctgag 1090

```

```

Oy 1080 CCAGGATACAGATGCAAGTNTGATGAGGAAGTCTGATGAGTCAAGATGATTAGT 1135
    |||
Db 1091 ccagggtacagatgcaagctcttgatgaggaactgatcgagctcaagaagacttgat 1146

```

RESULT 9

AAS58752
ID AAS58752 standard; cDNA; 691 BP.

AC AAS58752;

DT 13-FEB-2002 (first entry)

DE cDNA #1428 encoding portion of a human colon tumour protein.

KM Human; colon tumour protein; colon cancer; gene therapy; cytostatic; ss.

XX Homo sapiens.

PN W0200173027-A2.

PD 04-OCT-2001.

PF 22-MAR-2001; 2001WO-US09246.

PR 24-MAR-2000; 2000US-191597P.

PR 04-MAY-2000; 2000US-202024P.

PR 05-MAY-2000; 2000US-202189P.

PA (CORI-) CORIXA CORP.

PI Meagher MJ, Xu J, King GE;

DR WPI: 2001-611627/70.

PT New colon tumour proteins and related nucleic acid, useful for

PS treatment, prevention, diagnosis and monitoring of cancer.

XX Claim 4; Page 281; 299pp; English.

CC Th present invention relates to the isolation of novel cDNA sequences
CC encoding for at least an immunogenic portion of human colon tumour
CC proteins. The sequences of the invention are useful in pharmaceutical
CC compositions and vaccines for the prevention and treatment of cancers
CC such as colon cancer. They are also useful for the diagnosis and
CC monitoring of such cancers. Antibodies to the colon tumour proteins
CC and antigen presenting cells that express polynucleotides encoding
CC colon tumour proteins can be used to inhibit the development of
CC cancers. T-cells that react specifically with colon tumour proteins
CC are useful for removing tumour cells from samples (e.g. blood) and
CC for cancer treatment. The polynucleotide sequences are also useful in
CC gene therapy. AAS57325-AAS58880 represent the cDNA sequences of the
CC invention that encode for portions of human colon tumour proteins.

SQ Sequence 691 BP; 183 A; 168 C; 184 G; 140 T; 16 other;

Query Match 43.5%; Score 516.6; DB 23; Length 691;

Best Local Similarity 91.4%; Pred. No. 1.2e-136;

Matches 616; Conservative 0; Mismatches 46; Indels 12; Gaps 7;

```

Oy 200 CCGTGATGTGAGGAGCAATGGGCGCATTCACCGGATCCCTGAGCGTGGTTGAACATCC 259
    |||
Db 1 cctggatgtgagagcaatgggcgcatccacccgagctccgagcctggttgaacatcc 60
Oy 260 TCATCCCTGTGTTAGACCGGATCCGATATGTGACAGTCTCAAGAAATTTGTCATCAACG 319
    |||
Db 61 tcattccctgtgttagaccggatccgatatgtgcagagctccaagaaattgtcatcaacg 120
Oy 320 TCCCTGAGCAGTGGCGCTGTGACTCTCGACATGTAATCTGCAAAATGCATGGAGCTCTTT 379
    |||
Db 121 tgcctgagcagtggtgctgtgactctcgacatgtaatctgcaaaatgcattggagctctt 180

```

QY	380	ACCTGGCATATGATGACCCCTTACAAAGCAACCTCCGGTGGGAGACCCCTGAGATATCCG	439
Db	181	accgcgcatactatgaccccttaacaagccaagctccggtctgagagaccctcgatattgccc	240
QY	440	TCACCCAGCTAGTCTCAACAACAACATGAGATCAGAGCTCGCCAAACTCTCTNTGGACAAG	499
Db	241	tcacccagctagctccaacaacccaatgagatcagagctcggcnaactctctctgacaag	300
QY	500	TCCTTCGGGAAACGGGAGTCCCTGTAATCCGACATTTGGTGGATTCATCAACCAAGCTCCTG	559
Db	301	tcttcgcgggaacccggagatccctcgatgacacatctgagatgcataaccaaagctgctg	360
QY	560	ACTGCTGGGGTATCCCGCTGCCTCCTCCTGTTATGAGATCAA -GGATATCATCTGTGCCACCCCG	618
Db	361	actgctggggtatccgcgtgcctccgttatgagatcaagagatatcatctgycaccccgg	420
QY	619	GTTGAAAG -AGTCCTATGCAAGATGAGGTGGAGGCGAGAC -GGCGGAAACGGGGCCACATTC	676
Db	421	gttgaaagaagctctatgcacatgacagctgagagccagagccgggaacccggccacagctc	480
QY	677	TAGAGTCTGAGGGGACCCGAGCTGCGCCATCATATGTGGCGAAGG ----GAAGAACA	731
Db	481	tagagctctgaggggacccggaagctcgagatcatcaatctggtgcanaangggagaaacang	540
QY	732	GGCCCAAGTCTCTGGCCT -CCGAGACGAAAAAGGCTTAACAGA -TAAATAGGCACAG	788
Db	541	ggcccaagctctctgacctctcnaaacaataaagaatgaaacanaataaactccgacacagg	600
QY	789	AG -AGGCGATGTCAGTTCTTGCGCAGAGCCCAAGGCTTAAGGCTGAAGCTATTTCGAATCCTGG	847
Db	601	anaagcccatctgmgtttttgcnangccaaaggttaaacactgaancntttcgatcctg	660
QY	848	CTGACGAGCTCTACA	861
Db	661	gtgnacttttacc	674

RESULT 10
AAK93192
ID AAK93192 standard; cDNA; 567 BP.

DT 06-NOV-2001 (first entry)

Human cDNA clone representative sequence, SEQ ID NO: 1652

Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

OS Homo sapiens.

PN EP1130094-A2.

PD 05-SEP-2001.

PF 07-JUL-2000; 2000EP-0114089.

PR 08-JUL-1999; 99JP-0194486.

PR 02-MAY-2000; 2000JP-0183765.

PA (HELI-) HELIX RES INST.

PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

XX	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
XX	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

DR WPI; 2001-524255/58.

PT 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -

PS Example 11; SEQ ID NO 1652; 1380pp + sequence listing; English.

XX The invention relates to primers for synthesising full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5' - and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesising the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence was used as the
CC representative sequence from a human clone which was used in
CC homology searches to identify the clone.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
S0 Sequence 567 BP; 103 A; 162 C; 165 G; 125 T; 12 other:

Query Match	43.5%;	Score 516.4;	DB 22;	Length 567;
Best Local Similarity	97.0%;	Pred. No. 1.3e-136;		
Matches 550;	Conservative 0;	Mismatches 14;	Indels 3;	Gaps 3

29 30 TCGTTGTTCCGGAGGTCGCTGCGCGGTGGGAATGCTGGCGCGCGCGGGGCA 89

Db 1 tcgttggtccgaggtcgcctgcggcggtgggaatgtgcgcgcgcgcgcgc- ggggca 59

90 CTGGGGCCCTTTGCTGAGGGCTCTCTACTGGCTTCTGGCCGCGCTCCG-GCCGCGCT 14

Db 60 ctggggcccttctgtgagggtctctactgtctctctgagcggtccgagcgagcct 11

149 CCTCTGATTCGCCGGAACACACCGTGCTACTGTTCTGTCGCCGACGACGAGGCCCTGGGTGG 20

Db 120 cctctgattgcccgaaacacggtgtactgtctgtgccgacgagagcctggttg 17

209 TGGAGCGAATGGCCGATTCCACCGGATCCTGGAGCCTGGTTGAACATCCTCATCCCTG 260

Db 180 `tggaagcgaatgggccgattccaccggaatcctggagcctggttgaacatcctcatccctg` 23

269 TGTAGACCGGATCCGATATGTGCAGAGTCTCAAGGAATTGTCATCAACGTGCCCTGAGC 32

Db 240 tgttagaccgataccgatatgtgacagatcctcaaggaattgtcatcaacgtgcctgagc 255

329 AGTCGGCTGTGACTCTCGACAATGTAACTCTGCAATCGATGGAGTCCITTACCTGGCA 38

Db 300 agtcggtgtgactctcgacaatgttaactctgcnatcgatgagttccttacctgcgca 355

389 TCATGGACCCCTTACAAGGCAAGCTACGGTGTGGAGGACCCCTGAGTATGCCGTCACCCAGC 44

Db 360 tcatggaaccttacaaagcnagctacggtgtgagagaccctgagtatgcccgtcaccacgc 411

449 TAGCTCAAACCAACCATGAGATCAGAGCTCGGCAACTCTCTNTGGACCAAGTCTTCCGGG 50

Db 420 tagctcnaacaacccatganaatnaganctcgcgaactctctctgacaaagtctctccgg 47

509 AACGGAGTCCCTGAATGCCAGCATGTGGATGCCAT - CAACCAAGCTGCTGACTGCTGG 56

Db 480 aacgagagtcctgaatgccagcatgttngatgcatccaaccaagctgtgtaantgtctg 53

QY 568 GGATACCGGCTGCCCTNCGTTATGAGATC 594

Db 540 ggtatccgctgcccctccggttatnanaac 566

RESULT 11

AAH05113
ID AAH05115 standard: cDNA: 567 bp

XX
AC
AAH05175:

DT 26-JUN-2001 (first entry)

XX	Human cDNA clone (5'-primer)	SEQ ID NO:1950
DE		

Db 477 tagagctcgaaggagcccgagcgtcggccttaatgctgagcagaaggaagaacagagccc 536
Qy 737 AGATCCCTGGGCTCCGAGCAGCAAAAAGGCTGAACAGATTAATCAGCAGCAGAGAGCCCA 796
Db 537 caatcccgccctcgaagcaaaaaggctgacagaataaataaagcgcagagagaagccc 596
Qy 797 GTGCAGTTCTGCGAAGCCCAAGCTTAAGCTGAAGCTATTGCAATCTGCTGCAGCTC 856
Db 597 atcatcttgcgaagg-caagcctaagctgaaagctgcttcaaacctgcctgagcctt 655
Qy 857 TGACACAAATTAAT-GGAGATGCGAGCAGCTTCAC-TGACTGTGGCCGAGCAGTATGTCAG 914
Db 656 ttgacacacataatgagagcagccacctactgactgctgagcgtgagcagacatacagcag 715
Qy 915 GCGGTTCTCAAACT-GGCCAAGAGACTCAACACTATCTACTGCC 959
Db 716 cccggttccaactgagcgaagacatacaatcttctctatgagc 761

RESULT 14
AAK93925/c
ID AAK93925 standard; cDNA: 566 BP.

AAK93925;
06-NOV-2001 (first entry)

Human cDNA clone representative sequence, SEQ ID NO: 2385.

Human: full length cDNA: cDNA synthesis: oligo-capping; ss.

Homo sapiens.

EP1130094-A2.

05-SEP-2001.

07-JUL-2000; 2000EP-0114089.

08-JUL-1999; 99JP-0194486.

11-JAN-2000; 2000JP-0118774.

02-MAY-2000; 2000JP-0183765.

(HELI-) HELIX RES INST.

Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

WPI: 2001-524255/58.

830 primers useful for synthesizing full length cDNA clones and their

use in genetic manipulation -

Example 11: SEQ ID NO 2385; 1380bp + sequence listing: English.

The invention relates to primers for synthesizing full length cDNA clones. 830 cDNA molecules encoding a human protein have been isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have been determined. Primers for synthesizing the full length cDNA are useful for clarifying the function of the protein encoded by the cDNA. The full length clones were obtained by construction of full length enriched cDNA libraries that were synthesised by the oligo-capping method. The primers enable the production of the full length cDNA easily without any special methods. The present sequence was used as the representative sequence from a human clone which was used in

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in CD-ROM format directly from EPO.

Sequence 566 BP: 120 A; 152 C; 139 G; 155 T; 0 other;

Query Match 39.2%; Score 465.6; DB 22; Length 566;

Best Local Similarity 98.7%; Pred. No. 3.5e-122;
Matches 468; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 715 GCAGAAGGGAAGAAACAGGCCCCAGATCTGTGCTCCGAAGCAGAAAAGGCTGAACAGATA 774
Db 566 GCAGAGGAAGAAACAGGCCCCAGATCTGTGCTCCGAAGCAGAAAAGGCTGAACAGATA 507
Qy 775 AATCAGCAGCAGAGAGAGCCAGTGTGCAATCTTGTGCGAAGGCCCAAGGCTTAAGCTGAAGCT 834
Db 506 AATCAGCAGCAGAGAGAGCCAGTGTGCAATCTTGTGCGAAGGCCCAAGGCTTAAGCTGAAGCT 447
Qy 835 AATGGAATCTGGGCTGCGAGCTCTGACACATATATGAGATGAGCAGCTTCACTGACT 894
Db 446 AATGGAATCTGGGCTGCGAGCTCTGACACATATATGAGATGAGCAGCTTCACTGACT 387
Qy 895 GTGGCCGAGCAGATATGTCAGGCGCTTCTCAAACTGGCCAAAGAGACTCCACACTATCTCTA 954
Db 386 GTGGCCGAGCAGATATGTCAGGCGCTTCTCAAACTGGCCAAAGAGACTCCACACTATCTCTA 327
Qy 955 CTGCCCCCAACCCCTGGGCGATGTCAACAGCATGTGGCTCAAGGCCATGGGTATATGGA 1014
Db 326 CTGCCCCCAACCCCTGGGCGATGTCAACAGCATGTGGCTCAAGGCCATGGGTATATGGA 267
Qy 1015 GCGCTCACCAAGCCCGCAGTGCAGGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGA 1074
Db 266 GCGCTCACCAAGCCCGCAGTGCAGGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGA 207
Qy 1075 GATGTCCAGGGTACAGATGCAAGTMTGATGAGAACTTGATGAGTCAAGTGAATTAG 1134
Db 206 GATGTCCAGGGTACAGATGCAAGTMTGATGAGAACTTGATGAGTCAAGTGAATTAG 147
Qy 1135 TGGAGCTGGGCTTNGCCAGGGGAGCTGGGGACACAGAACGATTTTCTGATT 1188
Db 146 TGGAGCTGGGCTTNGCCAGGGGAGCTGGGGACACAGAACGATTTTCTGATT 93

RESULT 15
ABL05449
ID ABL05449 standard; cDNA: 1153 BP.

ABL05449;

26-MAR-2002 (first entry)

Drosophila melanogaster expressed polynucleotide SEQ ID NO 10829.

Drosophila: developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ss.

OS Drosophila melanogaster.

WO200171042-A2.

27-SEP-2001.

23-MAR-2001; 2001WO-US09231.

23-MAR-2000; 2000US-191637P.

11-JUL-2000; 2000US-0614150.

(PEKE) PE CORP NY.

Venter JC, Adams M, Li PMD, Myers EW;

WPI: 2001-656600/75.

P-PSDB; ABB61346.

New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

Claim 1; SEQ ID NO 10829; 21pp + Sequence Listing; English.

...

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 13:26:15 ; Search time 50.04 Seconds
(without alignments)
5831.587 Million cell updates/sec

Title: US-09-898-216-2
Perfect score: 1188
Sequence: 1 GCGTCTCGAGACGACCGCT.....GGAAGCAGATTTCGATT 1188

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2-6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2-6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2-6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2-6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2-6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2-6/ptodata/1/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1183	99.6	1188	1	US-08-781-562-2
2	48.8	4.1	7218	1	US-08-232-463-14
3	39.6	3.3	2488	1	US-08-279-270A-2
4	39.2	3.3	2557	4	US-08-464-954A-1
5	39	3.3	2204	1	US-08-221-817-12
6	39	3.3	2204	1	US-08-454-439-12
7	39	3.3	2204	5	PCT-US94-10487-12
8	39	3.3	2206	1	US-08-221-817-10
9	39	3.3	2206	1	US-08-454-439-10
10	39	3.3	2206	5	PCT-US94-10487-10
11	39	3.3	7791	2	US-08-149-097D-23
12	39	3.3	7791	3	US-08-949-386-23
13	39	3.3	7791	3	US-08-450-562-23
14	39	3.3	7791	4	US-08-984-709A-23
15	39	3.3	7808	4	US-08-149-097D-22
16	39	3.3	7808	3	US-08-949-386-22
17	39	3.3	7808	3	US-08-450-562-22
18	39	3.3	7808	4	US-08-984-709A-22
19	37.6	3.2	289	4	US-09-007-005-17
20	37.6	3.2	289	4	US-09-244-796-17
21	37.4	3.1	2848	4	US-08-464-954A-2
22	37.2	3.1	977	6	5215895-2
23	36.2	3.0	1983	1	US-08-221-817-21
24	36.2	3.0	1983	1	US-08-454-439-21
25	36.2	3.0	1983	5	PCT-US94-10487-21
26	36.2	3.0	3224	5	US-09-079-415-3
27	35.6	3.0	977	1	US-08-017-522A-1

28	35.6	3.0	1100	1	US-07-949-516A-1	Sequence 1, Appli
29	35.6	3.0	1100	2	US-08-814-459-1	Sequence 1, Appli
30	35.6	3.0	1100	3	US-09-122-525-1	Sequence 1, Appli
31	35.2	3.0	13842	4	US-09-105-537-30	Sequence 30, Appli
32	35.2	3.0	36778	4	US-09-105-537-5	Sequence 5, Appli
33	35.2	3.0	38506	3	US-09-320-878-19	Sequence 19, Appli
34	35	2.9	500	3	US-09-141-000-2	Sequence 2, Appli
35	34.8	2.9	2018	2	US-08-557-973-1	Sequence 1, Appli
36	33.8	2.8	2277	1	US-08-676-967-2	Sequence 2, Appli
37	33.8	2.8	2277	1	US-08-676-974-2	Sequence 2, Appli
38	33.8	2.8	2277	2	US-09-098-487-2	Sequence 2, Appli
39	33.4	2.8	8252	1	US-08-046-585-15	Sequence 15, Appli
40	33.4	2.8	8252	1	US-08-393-703-15	Sequence 2, Appli
41	33.4	2.8	8252	5	PCT-US93-11721-15	Sequence 15, Appli
42	33.2	2.8	2307	3	US-08-942-008-1	Sequence 1, Appli
43	32.8	2.8	1812	2	US-08-735-041A-3	Sequence 3, Appli
44	32.8	2.8	1812	3	US-09-190-476B-3	Sequence 3, Appli
45	32.8	2.8	1812	3	US-09-190-889A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-781-562-2
; Sequence 2, Application US/08781562
; Patent No. 5763589
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: NOVEL HUMAN MEMBRANE PROTEIN
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,562
FILING DATE: Herewith
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PR-0181 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1188 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE: Consensus
US-08-781-562-2
Query Match 99.6%; Score 1183; DB 1; Length 1188;
Best Local Similarity 100.0%; Pred. No. 0;

```
Matches 1188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Oy	1	GGTTCCTGGGAGNACCGGTCCGCCTCCTCCGTTGCTTCGGAGAGGATCGCTCGCGCGTG	60
Dp	1	GGGTTCCTGGGAGCACCGGCTCCGCTGCTCTCTGGTTCCGGAGGTCCCTCGGGGGTG	60
Oy	61	GAATTCGTGGCGGCGGCGGCGGCGCACTGGGGCCCTTTTGTCTGAGGGGCTCTACT	120
Dp	61	GAATTCGTGGCGGCGGCGGCGGCGGCGCACTGGGGCCCTTTTGTGAGGGGCTCTACT	120
Oy	121	GGCTTCCTGGCGGCGGCTCCGGGCCGCTCTCTCGATTTCGCCGAACCGTGGTACTG	180
Dp	121	GGCTTCCTGGCGGCGGCTCCGGGCCGCTCTCTCGATTTCGCCGAACCGTGGTACTG	180
Oy	181	TTCCTCCCGCAGAGAGAGCGCTGGTGTGTGAGACCAATGGCGCATTCACCGGATCCTG	240
Dp	181	TTCCTCCCGCAGAGAGAGCGCTGGTGTGTGAGACCAATGGCGCATTCACCGGATCCTG	240
Oy	241	GAGCTTGTTTTGAACATCTCTATCCCTGTGTTAGACCGGATCCGATTATGTGAGAGTTC	300
Dp	241	GAGCTTGTTTTGAACATCTCTATCCCTGTGTTAGACCGGATCCGATTATGTGAGAGTTC	300
Oy	301	AAGAAATTGTCAATCAACGTCGCTTGAGACAGTGGGCTGTACCTGTGACAATGTAACTGTG	360
Dp	301	AAGAAATTGTCAATCAACGTCGCTTGAGACAGTGGGCTGTACCTGTGACAATGTAACTGTG	360
Oy	361	CAAATCGATGGAGTCTTTTACCTTGCCTGCGCATCATGAGACCCTTACAAAGCAAGTACGGTGTG	420
Dp	361	CAAATCGATGGAGTCTTTTACCTTGCCTGCGCATCATGAGACCCTTACAAAGCAAGTACGGTGTG	420
Oy	421	GAGAACCCCTGAGTATGCGCGTCAACCCAGCTTAGCTCAAAACACCATAGATCAGAAGTGGC	480
Dp	421	GAGAACCCCTGAGTATGCGCGTCAACCCAGCTTAGCTCAAAACACCATAGATCAGAAGTGGC	480
Oy	481	AAATCTCTTMTGACAAAGTCTTCCGGAACGGGAGTCCCTTAATGSCAGATTTGTGAT	540
Dp	481	AAATCTCTTMTGACAAAGTCTTCCGGAACGGGAGTCCCTTAATGSCAGATTTGTGAT	540
Oy	541	GCCATCAACCAAGCTGCTGACTCTGGGGTATCCGCTGCTNCGTTATGAGATCAAGAT	600
Dp	541	GCCATCAACCAAGCTGCTGACTCTGGGGTATCCGCTGCTNCGTTATGAGATCAAGAT	600
Oy	601	ATTCATTTGCGCACCCCGGGGGAANAAGTCTATGACAGATCAGATGAGGAGGCGCGG	660
Dp	601	ATTCATTTGCGCACCCCGGGGGAANAAGTCTATGACAGATCAGATGAGGAGGCGCGG	660
Oy	661	AAACGGGCGCACAGTTCTTAGAGTCTGAGGGGACCCGAGATCGGCCATCAATGTGGCAGAA	720
Dp	661	AAACGGGCGCACAGTTCTTAGAGTCTGAGGGGACCCGAGATCGGCCATCAATGTGGCAGAA	720
Oy	721	GGGAAAGAAAGGCCCAAGATCCCTGGGCTCCGAAGCAAAAAGCGTTGAACAGATAAATAG	780
Dp	721	GGGAAAGAAAGGCCCAAGATCCCTGGGCTCCGAAGCAAAAAGCGTTGAACAGATAAATAG	780
Oy	781	GCACAGAGAGAGCGCAGTGCAGTTTCTGGGGAAGGCCAAGGCTAAAGCTGAAAGCTATTGCA	840
Dp	781	GCACAGAGAGAGCGCAGTGCAGTTTCTGGGGAAGGCCAAGGCTAAAGCTGAAAGCTATTGCA	840
Oy	841	ATCTGCGCTGACGCTCTGACACACAATAATGAGATGACAGAGCTTCACTACTGTGGCC	900
Dp	841	ATCTGCGCTGACGCTCTGACACACAATAATGAGATGACAGAGCTTCACTACTGTGGCC	900
Oy	901	GAGCAGTATGTCAAGCGGTTCTCCAATGAGGCAAGGAATCCACACATATCTTACTGGCC	960
Dp	901	GAGCAGTATGTCAAGCGGTTCTCCAATGAGGCAAGGAATCCACACATATCTTACTGGCC	960
Oy	961	TCCAAACCTTGCGATGTCAACAGCATGTTGGCTCAGGGCCATGGGTATATGAGACCCTC	1020
Dp	961	TCCAAACCTTGCGATGTCAACAGCATGTTGGCTCAGGGCCATGGGTATATGAGACCCTC	1020
Oy	1021	ACCAAACCCCCAGTGGCCAGGGAATCCAGACTCACTTCTCAATGTGGAGCAGCAGAGATGTC	1080
Dp	1021	ACCAAACCCCCAGTGGCCAGGGAATCCAGACTCACTTCTCAATGTGGAGCAGCAGAGATGTC	1080

Oy	1081	CAGGGTCACAGTGCACAATNTTGATTGAGGAACPTGATCGGTCAAAATAGTAGTAAGTGAGC	1140
Db	1081	CAGGGTCACACTGCCAATNTTGATTGAGGAACCTGATCGAGTCAAAATAGTAGTAAGTGAGC	1140
Oy	1141	TGGGCTTTGCCAGGAGACTCTGGGGACACAAGSAGACAGATTTTCTCTGATT	1188
Db	1141	TGGGCTTTGCCAGGAGACTCTGGGGACACAAGSAGACCAATTTTCTCTGATT	1188

RESULT 2
US-08-232-463-14/c

;; Sequence 14, Application US/08232463
; Patent No. 5670367

```

; GENERAL INFORMATION:
;
; APPLICANT: DORNER, F.

```

APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.

; TITLE OF INVENTION: RECOMBINANT
; NUMBER OF SEQUENCES: 52
;

CONFESSION ADDRESS:
ADDRESSEE: Foley & Lardner
1000 Broadway New York City 10003

CITY: Alexandria

COUNTRY: USA
7TP: 22313-0

COMPUTER READABLE FORM:
MEDIUM TYPE. F10000

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/232,463
FILING DATE:

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA
;

```

APPLCATION NUMBER: US/01/935,313
FILING DATE:

APPLICATION NUMBER: EP 91 114 300.0
FILING DATE: 26-AUG-1991

NAME: BENT, Stephen A.
REGISTRATION NUMBER: 30

REFERENCE NUMBER: 30472/114 IMM

TELEPHONE: (703) 836-93
TELEFAX: (703) 683-4109

TELEX: 899149
INFORMATION FOR SEQ ID NO: 14

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs

```

TYPE: nucleic acid
STRANDEDNESS: single

```

;          TOPOLOGY: linear
;          IMMEDIATE SOURCE:

```

CLONE: F
US-08-2332-463-14

Query Match 4.18; Score 48.8; DB 1; Length 7218;

Best Local Similarity 2.68; Pred. No. 0.00051;

Matches 8; Conservative 184; Mismatches 116; Indels 0; Gaps 0;

617 GGGTGAAGAGTCTATGCAGATCAGGTGGAGGCAGACGGCGGAACGGGCCACAGTTTC 676

b
1362 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1303

677 TAGAGTCTGAGGGGACCCGAGAGCTCGGCCATCAATGCTGGCAGAGGGGAGAGAAACAGGCCC 736

b
1302 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 124

Db 1377 GCGGCCCTAAGAGAGAGTCAAGCGGAGAGTGAACCCCGGGCTCTGTGAGACGG 1433

RESULT 5
US-08-221-817-12
Sequence 12, Application US/08221817
Patent No. 5532151
GENERAL INFORMATION:
APPLICANT: Chantry, David
APPLICANT: Gray, Patrick W.
APPLICANT: Hoeckstra, Merle F.
TITLE OF INVENTION: A No. 5532151el G Protein-Coupled Receptor
TITLE OF INVENTION: Kinase GRK6
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESSEE: Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/221,817
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/123,932
FILING DATE: 17 SEP 1993
ATTORNEY/AGENT INFORMATION:
NAME: No. 5532151and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31981
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2204 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 31..1758
US-08-221-817-12

Query Match 3.3%; Score 39; DB 1; Length 2204;
Best Local Similarity 60.6%; Pred. No. 0.17;
Matches 63; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 557 CTGACTGCTGGGGTATCCGCTGCTNCGTATGAGATCAAGATATCATGTGCCACCCC 616
DB 1127 CTGACTGCTGGGGTATCCGCTGCTNCGTATGAGATCAAGATATCATGTGCCACCCC 616

QY 617 GGCTGAAGAGTCTATGATCAGATGCAGTGGAGGAGCAGACGGCGG 660
DB 1187 AGCAGAGGAAGAAGATCAACGCGGAGAGGTGAGCGGCTG 1230

RESULT 6
US-08-454-439-12
Sequence 12, Application US/08454439
Patent No. 5591618
GENERAL INFORMATION:

APPLICANT: Chantry, David
APPLICANT: Gray, Patrick W.
APPLICANT: Hoeckstra, Merle F.
TITLE OF INVENTION: A No. 5591618el G Protein-Coupled Receptor
TITLE OF INVENTION: Kinase GRK6
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESSEE: Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,439
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/221,817
FILING DATE: 31-MAR-1994
APPLICATION NUMBER: 08/123,932
FILING DATE: 17 SEP 1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5591618and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31981
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2204 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 31..1758
US-08-454-439-12

Query Match 3.3%; Score 39; DB 1; Length 2204;
Best Local Similarity 60.6%; Pred. No. 0.17;
Matches 63; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 557 CTGACTGCTGGGGTATCCGCTGCTNCGTATGAGATCAAGATATCATGTGCCACCCC 616
DB 1127 CTGACTGCTGGGGTATCCGCTGCTNCGTATGAGATCAAGATATCATGTGCCACCCC 616

QY 617 GGCTGAAGAGTCTATGATCAGATGCAGTGGAGGAGCAGACGGCGG 660
DB 1187 AGCAGAGGAAGAAGATCAACGCGGAGAGGTGAGCGGCTG 1230

RESULT 7
PCT-US94-10487-12
Sequence 12, Application PC/TUS9410487
GENERAL INFORMATION:
APPLICANT: ICOS Corporation
TITLE OF INVENTION: A Novel G Protein-Coupled Receptor
TITLE OF INVENTION: Kinase GRK6
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &


```

; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/10487
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/221,817
; FILING DATE: 31 MAR 1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/123,932
; FILING DATE: 17 SEP 1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Noland, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31981
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2204 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31..1758
; PCT-US94-10487-12

Query Match          3.3%; Score 39; DB 5; Length 2204;
Best Local Similarity 60.6%; Pred. No. 0.17;
Matches 63; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 557 CTGACTGCTGGGGTATCCGCTCCTNCCTTATGAGATCAAGGATATCCATGTGCCACCCC 616
    ||||| |||| | ||||| | ||||| | ||||| | ||||| | ||||| |
DB 1127 CTGACTGGTGGCGGCTCGGCTCCTCTGTACGAGATGATCGACGGCCAGTGGCCCTTCC 1186
    ||||| |||| | ||||| | ||||| | ||||| | ||||| | ||||| |

QY 617 GGGTAAAGAGTCTATGCAGATGCAGGTGAGGCGAGCGCGCG 660
    || | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1187 AGCAGAGGAAGAAGATCAAGCGGAGAGGTGAGCGGCTG 1230
    || | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
US-08-221-817-10
; Sequence 10, Application US/08221817
; Patent No. 5532151
; GENERAL INFORMATION:
; APPLICANT: Chantry, David
; APPLICANT: Gray, Patrick W.
; APPLICANT: Hoekstra, Merle F.
; TITLE OF INVENTION: A No. 5532151el G Protein-Coupled Receptor
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSER: Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
```

```

; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/221,817
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/123,932
; FILING DATE: 17 SEP 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5532151and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31981
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31..1926
; US-08-221-817-10

Query Match          3.3%; Score 39; DB 1; Length 2206;
Best Local Similarity 60.6%; Pred. No. 0.17;
Matches 63; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 557 CTGACTGCTGGGGTATCCGCTCCTNCCTTATGAGATCAAGGATATCCATGTGCCACCCC 616
    ||||| |||| | ||||| | ||||| | ||||| | ||||| | ||||| |
DB 1124 CTGACTGGTGGCGGCTCGGCTCCTCTGTACGAGATGATCGACGGCCAGTGGCCCTTCC 1183
    ||||| |||| | ||||| | ||||| | ||||| | ||||| | ||||| |

QY 617 GGGTAAAGAGTCTATGCAGATGCAGGTGAGGCGAGCGCGCG 660
    || | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1184 AGCAGAGGAAGAAGATCAAGCGGAGAGGTGAGCGGCTG 1227
    || | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 9
US-08-454-439-10
; Sequence 10, Application US/08454439
; Patent No. 5591618
; GENERAL INFORMATION:
; APPLICANT: Chantry, David
; APPLICANT: Gray, Patrick W.
; APPLICANT: Hoekstra, Merle F.
; TITLE OF INVENTION: A No. 5591618el G Protein-Coupled Receptor
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSER: Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
```

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/454,439
;; FILING DATE: 30-MAY-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/221,817
;; FILING DATE: 31-MAR-1994
;; APPLICATION NUMBER: 08/123,932
;; FILING DATE: 17 SEP 1993
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: No. 5591618and, Greta E.
;; REGISTRATION NUMBER: 35,302
;; REFERENCE/DOCKET NUMBER: 31981
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (312) 474-6300
;; TELEFAX: (312) 474-0448
;; TELEX: 25-3856
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2206 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: CDNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 31..1926
US-08-454-439-10

Query Match 3.3%; Score 39; DB 1; Length 2206;
Best Local Similarity 60.6%; Pred. No. 0.17;
Matches 63; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

OY 557 CTGACTGCTGGGGGTATCCGCTGCCCTGCTATGAGATCAAGATATCCATGTGCCACCCC 616
DB 1124 CTGACTGGTGGCGCTGCGCTGCTCTCTGACGAGATGATGCGACGCCAGTGCCTTCC 1183
OY 617 GCGTGAAGAGCTCTATGCAGATGCAGTGCAGAGCGAGCGCGG 660
DB 1184 AGCAGAGAGAGAAGATCAAGCGGAGAGAGTGGAGCGGCTG 1227

RESULT 10
PCT-US94-10487-10
Sequence 10, Application PC/TUS9410487
GENERAL INFORMATION:
APPLICANT: ICOS Corporation
TITLE OF INVENTION: A Novel G Protein-Coupled Receptor
TITLE OF INVENTION: Kinase GRK6
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/10487
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/221,817
FILING DATE: 31 MAR 1994
CLASSIFICATION:

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/123,932
;; FILING DATE: 17 SEP 1993
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Noland, Greta E.
;; REGISTRATION NUMBER: 35,302
;; REFERENCE/DOCKET NUMBER: 27866/31981
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (312) 474-6300
;; TELEFAX: (312) 474-0448
;; TELEX: 25-3856
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2206 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: CDNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 31..1926
PCT-US94-10487-10

Query Match 3.3%; Score 39; DB 5; Length 2206;
Best Local Similarity 60.6%; Pred. No. 0.17;
Matches 63; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

OY 557 CTGACTGCTGGGGGTATCCGCTGCCCTGCTATGAGATCAAGATATCCATGTGCCACCCC 616
DB 1124 CTGACTGGTGGCGCTGCGCTGCTCTCTGACGAGATGATGCGACGCCAGTGCCTTCC 1183
OY 617 GCGTGAAGAGCTCTATGCAGATGCAGTGCAGAGCGAGCGCGG 660
DB 1184 AGCAGAGAGAGAAGATCAAGCGGAGAGAGTGGAGCGGCTG 1227

RESULT 11
US-08-149-097D-23
Sequence 23, Application US/08149097D
Patent No. 5874236
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/149,097D
FILING DATE: 05-NOV-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/105,536
FILING DATE: 11-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US92/06903

```

: FILING DATE: 14-AUG-1992
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 07/914,231
: FILING DATE: 13-JUL-1992
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 07/868,354
: FILING DATE: 10-APR-1992
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 07/745,206
: FILING DATE: 15-AUG-1991
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 07/620,250
: FILING DATE: 30-NOV-1990
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 07/482,384
: FILING DATE: 20-FEB-1990
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 07/603,751
: FILING DATE: 04-APR-1989
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: WO PCT/US89/01408
: FILING DATE: 04-APR-1989
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 07/176,899
: FILING DATE: 04-APR-1988
: ATTORNEY/AGENT INFORMATION:
: NAME: Seidman, Stephanie L.
: REGISTRATION NUMBER: 33,779
: REFERENCE/DOCKET NUMBER: 6362-55038
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 238-0999
: INFORMATION FOR SEQ ID NO: 23:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7791 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: unknown
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 240..7037
: OTHER INFORMATION: /product="Alpha1A-2 subunit of
: OTHER INFORMATION: human calcium channel"
US-08-149-097D-23

Query Match          3.3%; Score 39; DB 2; Length 7791;
Best Local Similarity 47.8%; Pred. No. 0.31;
Matches 111; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 564 CTGGGGTATCCGCTGCTTCGTTATGATGATCAAGATATCCATGTCACCCGGGTGAA 623
DB 2333 CTTTGGAGACTACACCTCTCTTAATGTGTTCTTGCCATCGCTGTGACAAATCTGGCCAA 2392
QY 624 AGAGTCTATGCAATGATGAGGTGAGGACAGAGCGCGGAAGCGGCCACAGTTCTAGAGTC 683
DB 2393 CGCCAGAGAGCTCACACAGGTGAGGCGGACGACGACGAAGGAAGAAGACGACCGCA 2452
QY 684 TGAGGGAGCCCGAGAGTGGCCATCAATGTGCGAGAGGGAAGAAACAGGCCCGAGATCCT 743
DB 2453 GAAACTTGGCTTACAGAAAGCCAGAGAGGTGCGAGAAAGTAGTCTCTCTGCGGGCCAA 2512
QY 744 GGCCTCCGAGAGAAAGGCTGAGACAGATTAATCAGCAGCAGAGAGAGGCC 795
DB 2513 CATGCTATAGCTGTGAAAGACCAAGAAATCAAAAAGCCAGCCAGCTCC 2564

RESULT 12
: Sequence 23, Application US/08949386
: Patent No. 6090623
: GENERAL INFORMATION:
```

```

: APPLICANT: Harpold, Michael
: APPLICANT: Ellis, Steven
: APPLICANT: Williams, Mark
: APPLICANT: McCue, Ann
: APPLICANT: Gillespie, Alison
: TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
: NUMBER OF SEQUENCES: 38
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Brown, Martin, Haller & McClain
: STREET: 1660 Union Street
: CITY: San Diego
: STATE: California
: COUNTRY: US
: ZIP: 92101
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/949,386
: FILING DATE:
: CLASSIFICATION:
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US/08/290,012
: FILING DATE: 11-AUG-1994
: APPLICATION NUMBER: 08/149,097
: FILING DATE: 5-NOV-1993
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: 08/105,536
: FILING DATE: 11-AUG-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Seidman, Stephanie L.
: REGISTRATION NUMBER: 33,779
: REFERENCE/DOCKET NUMBER: 519808
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 238-0999
: INFORMATION FOR SEQ ID NO: 23:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7791 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 237..7037
: OTHER INFORMATION: /standard_name="Alpha-1A-2"
US-08-949-386-23

Query Match          3.3%; Score 39; DB 3; Length 7791;
Best Local Similarity 47.8%; Pred. No. 0.31;
Matches 111; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 564 CTGGGGTATCCGCTGCTTCGTTATGATGATCAAGATATCCATGTCACCCGGGTGAA 623
DB 2333 CTTTGGAGACTACACCTCTCTTAATGTGTTCTTGCCATCGCTGTGACAAATCTGGCCAA 2392
QY 624 AGAGTCTATGCAATGATGAGGTGAGGACAGAGCGCGGAAGCGGCCACAGTTCTAGAGTC 683
DB 2393 CGCCAGAGAGCTCACACAGGTGAGGCGGACGACGACGAAGGAAGAAGACGACCGCA 2452
QY 684 TGAGGGAGCCCGAGAGTGGCCATCAATGTGCGAGAGGGAAGAAACAGGCCCGAGATCCT 743
DB 2453 GAAACTTGGCTTACAGAAAGCCAGAGAGGTGCGAGAAAGTAGTCTCTCTGCGGGCCAA 2512
QY 744 GGCCTCCGAGAGAAAGGCTGAGACAGATTAATCAGCAGCAGAGAGAGGCC 795
DB 2513 CATGCTATAGCTGTGAAAGACCAAGAAATCAAAAAGCCAGCCAGCTCC 2564
```

RESULT 13
US-08-450-562-23
Sequence 23, Application US/08450562
Patent No. 6096514
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: McCue, Ann
APPLICANT: Gillespie, Allison
APPLICANT: Feldman, Daniel
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: US
ZIP: 92101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,562
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/404,950
FILING DATE: 13-MAR-1995
APPLICATION NUMBER: 08/336,257
FILING DATE: 7-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/314,083
FILING DATE: 28-SEPT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/311,363
FILING DATE: 23-SEPT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/290,012
FILING DATE: 11-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/223,305
FILING DATE: 4-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/193,078
FILING DATE: 07-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/149,097
FILING DATE: 5-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/105,536
FILING DATE: 11-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/914,231
FILING DATE: 13-JULY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/868,354
FILING DATE: 10-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/06903
FILING DATE: 14-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/620,250
FILING DATE: 30-NOV-1990

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/603,751
FILING DATE: 08-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/482,384
FILING DATE: 02-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US89/01408
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/176,899
FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-519812
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 238-0999
TELEFAX: (619) 238-0062
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 7791 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 237..7037
OTHER INFORMATION: /standard_name="Alpha-1A-2"
US-08-450-562-23
Query Match: 3.3%; Score 39; DB 3; Length 7791;
Best Local Similarity 47.8%; Pred. No. 0.31; Indels 0; Gaps 0;
Matches 111; Conservative 0; Mismatches 121;
QY 564 CTGGGGTATCCGCTCCCTGCTTATGATCAAGATATCCATGCCACCCGGGTGA 623
DB 2333 CTTTGGGAATCAGCCCTCCGTAATGTGTTTGCCATCCGTGGCAATCTGGCCAA 2392
QY 624 AGAGTCTATGCAGATGCAGGTGGAGCGGCGGGAAGCCACAGTTCTAGAGTC 683
DB 2393 CGCCCGAGAGCTCACCAGGTGGAGCGGAGAGGAAGGAAGCAGCAGCA 2452
QY 684 TGAGGGAGCCGAGAGCTGCGCCATCATGTGGCAGAGGAAGAAACAGCCAGATCCT 743
DB 2453 GAAACTTGCCCTACAGAAAGCCAGAGAGGTGGCAGAAAGTGAAGTCTCTGTGCGGCCAA 2512
QY 744 GGCCTCGAAGCAGAAAGGCTGAACAGATTAATCAGGCGAGCAGAGAGGCC 795
DB 2513 CATGCTATAGCTGTGAAGACACAGAAATCAAAAGCCAGCAAGTGC 2564
RESULT 14
US-08-984-709A-23
Sequence 23, Application US/08984709A
Patent No. 6320032
GENERAL INFORMATION:
APPLICANT: Williams, Mark E.
APPLICANT: Stauderman, Kenneth A.
APPLICANT: Harpold, Michael M.
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Heller Ehrman White & McCulliffe
STREET: 4250 Executive Square, Suite 700
CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:

Db 2453 GAACTTGCCCTACAGAAAGCCAGAGAGTGCGACAAGTGAGTCCTCTGTCCGGGCCAA 2512
Oy 744 GGCCCTCCGAGCAGAAAAGCCTGAACAGATAAATCAGGCAGCAGAGAGGCC 795
Db 2513 CATGCTATAGCTGTGAAAGAGACAAGAGATCAAAAAGCCAGCCAAAGTCC 2564

Search completed: September 22, 2002, 15:22:50
Job time: 6995 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 14:20:45 : Search time 2868.76 Seconds
(without alignments)
8959.362 Million cell updates/sec

Title: US-09-898-216-2
1188
Sequence: 1 GGCTTCGGAGACNACCGCT.....GGAGCAGATTTCCTGATT 1188

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_NA_Main:*

1: /cgn2_6/ptodata/1/pna/US06005.COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06006.COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07007.COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US08008.COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US08009.COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US08010.COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US08011.COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US08012.COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US08013.COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US08014.COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US08015.COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US08016.COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US08017.COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US08018.COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US08019.COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US08020.COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US08021.COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US08022.COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US08023.COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US08024.COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US08025.COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US08026.COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US08027.COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US08028.COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US08029.COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US08030.COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US08031.COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US08032.COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US08033.COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US08034.COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US08035.COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US08036.COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US08037.COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US08038.COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US08039.COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US08040.COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US08041.COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US08042.COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US08043.COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US08044.COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US08045.COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US08046.COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US08047.COMB.seq.*

44: /cgn2_6/ptodata/1/pna/US06005.COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US06006.COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US06007.COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US06008.COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US06009.COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US06010.COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US06011.COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US06012.COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US06013.COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US06014.COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US06015.COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US06016.COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US06017.COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US06018.COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US06019.COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US06020.COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US06021.COMB.seq.*
61: /cgn2_6/ptodata/1/pna/US06022.COMB.seq.*
62: /cgn2_6/ptodata/1/pna/US06023.COMB.seq.*
63: /cgn2_6/ptodata/1/pna/US06024.COMB.seq.*
64: /cgn2_6/ptodata/1/pna/US06025.COMB.seq.*
65: /cgn2_6/ptodata/1/pna/US06026.COMB.seq.*
66: /cgn2_6/ptodata/1/pna/US06027.COMB.seq.*
67: /cgn2_6/ptodata/1/pna/US06028.COMB.seq.*
68: /cgn2_6/ptodata/1/pna/US06029.COMB.seq.*
69: /cgn2_6/ptodata/1/pna/US06030.COMB.seq.*
70: /cgn2_6/ptodata/1/pna/US06031.COMB.seq.*
71: /cgn2_6/ptodata/1/pna/US06032.COMB.seq.*
72: /cgn2_6/ptodata/1/pna/US06033.COMB.seq.*
73: /cgn2_6/ptodata/1/pna/US06034.COMB.seq.*
74: /cgn2_6/ptodata/1/pna/US06035.COMB.seq.*
75: /cgn2_6/ptodata/1/pna/US06036.COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1166	98.1	1322	1 PCT-US98-12125-102	Sequence 102, App
2	1166	98.1	1322	1 US-09-209-462B-106	Sequence 106, App
3	1160	97.6	1398	1 PCT-US01-03800A-1001	Sequence 1001, App
4	1160	97.6	1398	1 PCT-US01-04098A-2212	Sequence 2212, App
5	1160	97.6	1398	1 US-09-471-275-865	Sequence 865, App
6	1160	97.6	1398	1 US-09-496-914A-8410	Sequence 8410, App
7	1160	97.6	1398	1 US-09-560-875A-8410	Sequence 8410, App
8	1155	97.2	1384	1 US-09-338-467-941	Sequence 941, App
9	1155	97.2	1384	1 US-09-644-869-8024	Sequence 8024, App
10	1155	97.2	1384	1 US-09-649-162-7309	Sequence 7309, App
11	1155	97.2	1384	1 US-09-652-109-7925	Sequence 7925, App
12	1155	97.2	1384	1 US-09-652-127-7358	Sequence 7358, App
13	1155	97.2	1384	1 US-09-652-816-7419	Sequence 7419, App
14	1155	97.2	1384	1 US-09-652-913-8809	Sequence 8809, App
15	1155	97.2	1384	1 US-09-652-914-7997	Sequence 7997, App
16	1155	97.2	1384	1 US-09-808-384-941	Sequence 941, App
17	1144.4	96.3	1337	1 PCT-US98-12125-45	Sequence 45, App
18	1144.4	96.3	1337	1 US-09-209-462B-45	Sequence 45, App
19	1144	96.3	1416	1 PCT-US01-04098A-244	Sequence 244, App
20	1144	96.3	1416	1 US-09-598-075-249	Sequence 249, App
21	1144	96.3	1416	1 US-09-598-075A-249	Sequence 249, App
22	1105.4	93.0	1429	1 US-60-172-373-10483	Sequence 10483, App
23	1104.8	93.0	9098	1 PCT-US01-04942A-34	Sequence 34, App
24	1104.8	93.0	9098	1 US-09-596-193A-34	Sequence 34, App
25	1086.4	91.4	1234	1 US-60-164-285-3468	Sequence 3468, App
26	1072	90.2	1308	1 US-09-205-070-13379	Sequence 13379, App
27	1072	90.2	1308	1 US-09-340-633-13379	Sequence 13379, App
28	1072	90.2	1308	1 US-09-898-888-13379	Sequence 13379, App
29	1067.6	89.9	1501	1 US-09-898-888A-13379	Sequence 13379, App
30	1067.6	89.9	1501	1 US-09-298-733-35	Sequence 35, App
31	1067.6	89.9	1501	1 US-09-298-733A-35	Sequence 35, App

32	1067.6	89.9	1501	29	US-09-723-594-35	Sequence 35, Appl
33	1067.6	89.9	1501	29	US-09-724-497-35	Sequence 35, Appl
c 34	1057.8	89.0	1309	17	US-09-359-922-11863	Sequence 11863, Appl
c 35	1057.8	89.0	1309	17	US-09-359-922-11863	Sequence 11863, Appl
c 36	1042.8	87.8	9278	1	PCT-US01-08656-5097	Sequence 5097, Ap
37	863.4	73.2	1237	75	US-60-360-207-18511	Sequence 18511, A
38	728.4	61.3	732	33	US-09-878-134-349	Sequence 349, Appl
39	728.4	61.3	732	37	US-10-033-356-349	Sequence 349, Appl
40	683.6	57.5	1329	57	US-60-185-213-1569	Sequence 1569, Ap
41	674.8	56.8	805	17	US-09-393-720-18015	Sequence 18015, A
42	674.8	56.8	805	34	US-09-921-178-18015	Sequence 18015, A
c 43	618.6	45.0	522	36	US-09-998-998-1161	Sequence 1161, Ap
c 44	546.4	46.0	550	1	PCT-US02-02870-2995	Sequence 2995, Ap
c 45	546.4	46.0	550	37	US-10-066-543-2995	Sequence 2995, Ap

ALIGNMENTS

RESULT 1
PCT-US98-12125-102

```

: Sequence 102, Application PC/TUS98121255
: GENERAL INFORMATION:

```

Query Match	98.1%	Score 1166	DB 1	Length 1322
Best Local Similarity	99.2%	Pred. No. 8.1e-280		
Matches 1177: Conservative	1	Mismatches 7	Indels 1	Gaps 1

3 CTCTGAGGAGNACCCGCTCCGCTGCTGTCGTTGGTTTCGAGAGCTCGATGGCGGCGTGGGA 62
OY
Db 1 CTCTGAGGAGCACCGCTCCGCTGCTGCTGTTGGTTTCGAGAGCTCGATGGCGGCGTGGGA 60
OY 63 AATGCTGGCGCGCGCGCGCGCGGCGCACATGGGGCCCTTTTTCCTAGAGGGCTCTCTACTGG 1222
61 AATGCTGGCGCGCGCGCGCGCG -GNGGCACATGGGCGCCCTTTTTCCTAGAGGGCTCTCTACTGG 119b

QY	123	CTTGTGGCGCGGCTCGCGCGCGGCGCTCTCTGGATTTCGCCCGGAAACACCGTGGACGTGTT	182
Db	120	CTTCTGGCGCGGCTCGCGCGCGCTCTCTGGATTTCGCCCGGAAACACCGTGGACGTGTT	179
QY	183	CGTCCCGCAGCAGGAGGCGCTGGGTGGTGGAGGGAATGGGCGCGATTTCACCGCGATCTGGGA	242
Db	180	CGTCCCGCAGCAGGAGGCGCTGGGTGGTGGAGGGAATGGGCGCGATTTCACCGCGATCTGGGA	239
QY	243	GCCTGTGTTGAACATCTCATTCCTGTGTTTACACGGGATTCGATATTGTGCAGAGTCTCA	302
Db	240	GCCTGTGTTGAACATCTCATTCCTGTGTTTACACGGGATTCGATATTGTGCAGAGTCTCA	299
QY	303	GGAATTTGTATCAACCTGGCTGAGCAGTCCGGGTGATCTTCGACATATGTAACCTGCA	362
Db	300	GGAATTTGTATCAACCTGGCTGAGCAGTCCGGGTGATCTTCGACATATGTAACCTGCA	359
QY	363	AATGATGGAGTCTCTTACCTGGCATCATGACCCCTTACAAGGACGTAACGGTGTGGA	422
Db	360	AATGATGGAGTCTCTTACCTGGCATCATGACCCCTTACAAGGACGTAACGGTGTGGA	419
QY	423	GGACCCCTGAGTATCCCTCTACACCAAGTACGTCAAAACAATGACATGACAGTCCGGCA	482
Db	420	GGACCCCTGAGTATCCCTCTACACCAAGTACGTCAAAACAATGACATGACAGTCCGGCA	479
QY	483	ACTCTCTTTGGACAAATCTTCCGGGGAACGGGAGTCCCTGAAATGGCAGCATTTGTGGATGC	542
Db	480	ACTCTCTCTGGACAAATCTTCCGGGGAACGGGAGTCCCTGAAATGGCAGCATTTGTGGATGC	539
QY	543	CATCAACCAAGCTCTGCTAGCTGTGGGATTCGGCTCCCTTCGTTATGAGATCAAGATAT	602
Db	540	CATCAACCAAGCTCTGCTAGCTGTGGGATTCGGCTCCCTTCGTTATGAGATCAAGATAT	599
QY	603	CGATGTGCAACCCCGGGGTGAAAGAGTCTATGCGAGTACAGGTGGAGGCAAGCGCGCGAA	662
Db	600	CGATGTGCAACCCCGGGGTGAAAGAGTCTATGCGAGTACAGGTGGAGGCAAGCGCGCGAA	659
QY	663	ACGGGCCACAGTTCTAGAGTCTGAGGGGGAACCCGAGAAGTGGCCATCAATTGTGGCAGAAG	722
Db	660	ACGGGCCACAGTTCTAGAGTCTGAGGGGGAACCCGAGAAGTGGCCATCAATTGTGGCAGAAG	719
QY	723	GAAGAANAGGCCCAAGATNCCGTGGCCCTCCGAGCAGAAAGGCTGAAACAGATTAATCAAGC	782
Db	720	GAAGAANAGGCCCAAGATNCCGTGGCCCTCCGAGCAGAAAGGCTGAAACAGATTAATCAAGC	779
QY	783	AGCAGAGAGGCCAGTCCAGTTCTGGCGGAAGGCCAAGGCTAAAGCTGAAGCTATTTCGAT	842
Db	780	AGCAGAGAGGCCAGTCCAGTTCTGGCGGAAGGCCAAGGCTAAAGCTGAAGCTATTTCGAT	839
QY	843	CCTGGCTCAACCTTGACACAACTAATGAGATGCAGCAGCTTCACTGACTGTGGCGGA	902
Db	840	CCTGGCTCAACCTTGACACAACTAATGAGATGCAGCAGCTTCACTGACTGTGGCGGA	899
QY	903	GCAGTATGTACGCCGTTTCTCAAACTGGCGCAAGAGCTCAACACATATCTTACTGCCCTC	962
Db	900	GCAGTATGTACGCCGTTTCTCAAACTGGCGCAAGAGCTCAACACATATCTTACTGCCCTC	959
QY	963	CAACCTTGGCGATGTCAACAGCATGTGTGCTCAGGCCATATGGGTGTATATGAGCCCTCAC	1022
Db	960	CAACCTTGGCGATGTCAACAGCATGTGTGCTCAGGCCATATGGGTGTATATGAGCCCTCAC	1019
QY	1023	CAAAACCCCACTGGCAGGAGCTCCAGACATCTCTCAATGGGAGACAGCAGATGTGCA	1082
Db	1020	CAAAACCCCACTGGCAGGAGCTCCAGACATCTCTCAATGGGAGACAGCAGATGTGCA	1079
QY	1083	GGGTACAGATCAAGTATTGTATGAGAACTTGATCGAGTCAAGATAGATTACTGAGAGCTG	1142
Db	1080	GGGTACAGATCAAGTATTGTATGAGAACTTGATCGAGTCAAGATAGATTACTGAGAGCTG	1139
QY	1143	GGCTTNGCCAGGAGTCTGGGGACAGAGGAAGCATTTTCTCTGATT	1188
Db	1140	GGCTTNGCCAGGAGTCTGGGGACAGAGGAAGCATTTTCTCTGATT	1185


```

1 RESULT 2
2 US-09-209-462B-106
3 Sequence 106, Application US/092094462B
4 GENERAL INFORMATION:
5 APPLICANT: Rosen et al.
6 TITLE OF INVENTION: 86 Human Secreted Proteins
7 FILE REFERENCE: P2008P1
8 CURRENT APPLICATION NUMBER: US/09/209,462B
9 CURRENT FILING DATE: 1998-12-11
10 PRIOR APPLICATION NUMBER: PCT/US98/12125
11 PRIOR FILING DATE: 1998-06-11
12 PRIOR APPLICATION NUMBER: 60/049,547
13 PRIOR FILING DATE: 1997-06-13
14 PRIOR APPLICATION NUMBER: 60/049,548
15 PRIOR FILING DATE: 1997-06-13
16 PRIOR APPLICATION NUMBER: 60/049,549
17 PRIOR FILING DATE: 1997-06-13
18 PRIOR APPLICATION NUMBER: 60/049,550
19 PRIOR FILING DATE: 1997-06-13
20 PRIOR APPLICATION NUMBER: 60/049,556
21 PRIOR FILING DATE: 1997-06-13
22 PRIOR APPLICATION NUMBER: 60/049,606
23 PRIOR FILING DATE: 1997-06-13
24 PRIOR APPLICATION NUMBER: 60/049,607
25 PRIOR FILING DATE: 1997-06-13
26 PRIOR APPLICATION NUMBER: 60/049,608
27 PRIOR FILING DATE: 1997-06-13
28 PRIOR APPLICATION NUMBER: 60/049,609
29 PRIOR FILING DATE: 1997-06-13
30 PRIOR APPLICATION NUMBER: 60/049,610
31 PRIOR FILING DATE: 1997-06-13
32 PRIOR APPLICATION NUMBER: 60/049,611
33 PRIOR FILING DATE: 1997-06-13
34 PRIOR APPLICATION NUMBER: 60/050,901
35 PRIOR FILING DATE: 1997-06-13
36 PRIOR APPLICATION NUMBER: 60/052,989
37 PRIOR FILING DATE: 1997-06-13
38 PRIOR APPLICATION NUMBER: 60/051,919
39 PRIOR FILING DATE: 1997-07-08
40 PRIOR APPLICATION NUMBER: 60/055,984
41 PRIOR FILING DATE: 1997-08-18
42 PRIOR APPLICATION NUMBER: 60/058,665
43 PRIOR FILING DATE: 1997-09-12
44 PRIOR APPLICATION NUMBER: 60/058,668
45 PRIOR FILING DATE: 1997-09-12
46 PRIOR APPLICATION NUMBER: 60/058,669
47 PRIOR FILING DATE: 1997-09-12
48 PRIOR APPLICATION NUMBER: 60/058,750
49 PRIOR FILING DATE: 1997-09-12
50 PRIOR APPLICATION NUMBER: 60/058,971
51 PRIOR FILING DATE: 1997-09-12
52 PRIOR APPLICATION NUMBER: 60/058,972
53 PRIOR FILING DATE: 1997-09-12
54 PRIOR APPLICATION NUMBER: 60/058,975
55 PRIOR FILING DATE: 1997-09-12
56 PRIOR APPLICATION NUMBER: 60/060,834
57 PRIOR FILING DATE: 1997-10-02
58 PRIOR APPLICATION NUMBER: 60/060,841
59 PRIOR FILING DATE: 1997-10-02
60 PRIOR APPLICATION NUMBER: 60/060,844
61 PRIOR FILING DATE: 1997-10-02
62 PRIOR APPLICATION NUMBER: 60/060,865
63 PRIOR FILING DATE: 1997-10-02
64 PRIOR APPLICATION NUMBER: 60/061,059
65 PRIOR FILING DATE: 1997-10-02
66 PRIOR APPLICATION NUMBER: 60/061,060
67 PRIOR FILING DATE: 1997-10-02
68 NUMBER OF SEQ ID NOS: 737
69 SOFTWARE: PatentIn Ver. 2.0
70 SEQ ID NO 106
71 LENGTH: 1322
72 TYPE: DNA
73 ORGANISM: Homo sapiens

```

[illegible]

Db	780	agcagcgagagccagcgccgctcttcgcgaaggccaagctaaagctgaagctatcttgaat	839
QY	843	CTTGCTGCAGCTCTGCACACACATATATGAGATGCAGCAGCTTGACTGACTGTGGCGA	902
Db	840	ccctgcgcgcagcctctgacaacaacataatgagatgcagcagcttcaatgactctgcgcga	899
QY	903	GCAATATGTGAGGGCGTTCCTCCAAACTGGCCAAAGGACGTCCAAACAGTATATCATTACGCCCTC	962
Db	900	gcacatagctcaagcgctctctcccaactcgcgcgaagcctccacaacatcatccactcgcctc	959
QY	963	CAACCCCTGGCGATGTCCACCAGCATGTGTGGCTCAGGCCATGGGTGATATGAGCCCTCAC	1022
Db	960	caaccctcgcgcatgctcaaccagcactgttgctcgaagccatggtctgatcttgagccctcac	1019
QY	1023	CAAGCCCGCAGTGCACGGGACTCCACATCTCCTCCATGGGAGCAGCAGAGATGTCCA	1082
Db	1020	caaaagcccccagctgcacagggaccccaacatccactccagctggagacagacagagatgtcca	1079
QY	1083	GGGTACAGATATCCAAAGTTTTCATGAGAGACTTGATCGATGCAAGATGAGCTTATGATGAGCTG	1142
Db	1080	gggtacagatgcgaagctcttgatgaggaacttgatcgagctcaagagagagcttagtgagagctg	1139
QY	1143	GGCTTNNCCACAGAGCTGTGGGGACCAAGAACGACGATTTTCCATGATG	1188
Db	1140	ggctctggcccaaggagctctctggggaacaaggacagatcttccctgaat	1185

RESULT 3

```

PCT-US01-03800A-1001/C
: Sequence 1001, Application PC/TUS0103800A
: GENERAL INFORMATION:
: APPLICANT: Hyaseg, Inc.
: APPLICANT: Tang, Y., Tom et al
: TITLE OR INVENTION: Novel Nucleic Acids and Polypeptides
: FILE REFERENCE: 21272-030
: CURRENT APPLICATION NUMBER: PCT/US01/03800A
: PRIOR FILING DATE: 2001-02-05
: PRIOR APPLICATION NUMBER: 09/560, 875
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: 09/496, 914
: PRIOR FILING DATE: 2000-02-03
: NUMBER OF SEQ ID NOS: 2700
: SOFTWARE: Custom
: SEQ ID NO 1001
: LENGTH: 1398
: TYPE: DNA
: ORGANISM: Homo sapiens
PCT-US01-03800A-1001

```

Query Match 97.6%; Score 1160; DB 1; Length 1398;
Best Local Similarity 99.4%; Pred. No. 2.6e-278;
Matches 1182; Conservative 0; Mismatches 5; Indels 2; Gaps 2

Qy	1	GGCTTTGGGAGGCAACCGCTCCGCTGCTCGTGTGTTGGTCCGAGAGGCGCGCTGGCGGGTGG	60
Db	1393	GGCTTTCGGAGACGACCGGCTCCGCTCTCTCTGTGGTTCCGGAGGCGCGCTGGCGGGTGG	1334
Qy	61	GAATGCTGGCGCGCGCGCGCGGGGGGACATGGGGCCCTTTTGGCTGAGGGGCTCTTACT	120
Db	1333	GAATGCTGGCGCGCGCGCGGGGCGC-GGGGCACTGGGGCCCTTTTGGCTGAGGGGCTCTTACT	1275
Qy	121	GGCTTTCGGCGCGCGCTCCG-GCGCGCGCTCTCTTGATTGGCCCGGAAACACGCTGTACT	179
Db	1274	GGCTTTCGGCGCGCGCTCCGCGCGCGCTCTCTTGATTGGCCCGGAAACACGCTGTACT	1215
Qy	180	GTTTCGTGGCGGAGGAGGCGCTGGGTGGTGGAGGCAATGGGCGCGATTCACCGGATCT	239
Db	1214	GTTTCGTGGCGGAGGAGGCGCTGGGTGGTGGAGGCAATGGGCGCGATTCACCGGATCT	1155
Qy	240	GGAGCTGTGTTGAACATCTCATCCCTGTGTAGACCGGATCCGATATGTGACAGTCT	299

Db	1154	GGAGCCTGGTTTGAAACATCCTCATCCCTGTTGTTAGACCGGATCCGATATGTCGAGAGTCT	1095
OY	300	CAAGAAATTTGTCTATCAAGCTGCTCTAGACAGATGCGGTGTGACTCTTCGACATGTTAACTCT	359
Db	1094	CAAGAAATTTGTCTATCAACCTGTCTAGAGAGTCGCGGTGTGACTCTTCGACATGTTAACTCT	1035
OY	360	GCAATATGATGAGATCCTTTTACCTGTGGCATCATGACACCTTTCAAGGGACAGTCAGGTCGT	419
Db	1034	GCAATATGATGAGATCCTTTTACTGTGGCATCATGAGACCTTTCAAGGGACAGTCAGGTCGT	975
OY	420	GGAGAGCCCTGAGTATGCCGTCAACCCAGCTAGCTCAAAACCAATGAGATCAGAGCTCGG	479
Db	974	GGAGAGCCCTGAGTATGCCGTCAACCCAGCTAGCTCAAAACCAATGAGATCAGAGCTCGG	915
OY	480	CAAACTCTGTTGGACAAATCTTCCGGGAGGGGAGTCCCTGATATGCGAGATTGTGGA	539
Db	914	CAAACTCTCTGTGACAAATCTTCCGGGAGGGGAGTCCCTGATATGCGAGATTGTGGA	855
OY	540	TGCCATCAACCAAGCTTCTACTGCTGTGGGTATCCGACTCCCTGCTTATGAGATCAAGA	599
Db	854	TGCCATCAACCAAGCTTCTACTGCTGTGGGTATCCGACTCCCTGCTTATGAGATCAAGA	795
OY	600	TATCATATGTGCAACCCCGGGGTGAAAGAAGTATAGCAATGACAGGTGGAGGCAAGCCGCG	659
Db	794	TATCATATGTGCAACCCCGGGGTGAAAGAAGTATAGCAATGACAGGTGGAGGCAAGCCGCG	735
OY	660	GAACGGGGCCCAAGTTTCTAAGATCTGAGGGGAGCCGAGAGTGGGCGCATCAATGTGGCAGA	719
Db	734	GAACGGGGCCCAAGTTTCTAAGATCTGAGGGGAGCCGAGAGTGGGCGCATCAATGTGGCAGA	675
OY	720	AGGGAAGAAACAGGCCCAAGATCCTGGGCTTCGGAAGCAAAAAGCTGAAACAGATATAATCA	779
Db	674	AGGGAAGAAACAGGCCCAAGATCCTGGGCTTCGGAAGCAAAAAGCTGAAACAGATATAATCA	615
OY	780	GGCAGCAGAGAGAGCCAGTGCAGTTCGTGGCGAAGGCCCAAGGCTTAAAGCTGAAAGCTATTGCG	839
Db	614	GGCAGCAGAGAGAGGCCAGTGCAGTTCGTGGCGAAGGCCCAAGGCTTAAAGCTGAAAGCTATTGCG	555
OY	840	AATCCTGGCTGACGCTCTGACACACATTAATGAGATGCAAGCACTTCACTGACTGTGGC	899
Db	554	AATCCTGGCTGACACTCTGACACACATTAATGAGATGCAAGCACTTCACTGACTGTGGC	495
OY	900	CGAGAGATATGTACGCCGTTTCTCCAAATCGGGCAGAGACTCCAAACATATCTCTACTGCC	959
Db	434	CGAGAGATATGTACGCCGTTTCTCCAAATCGGGCAGAGACTCCAAACATATCTCTACTGCC	435
OY	960	CTTCAACCTTGGCGATGTACACAGCATGTGTGCTCAGGCCATGCGTGTATATGAGACCTT	1019
Db	434	CTTCAACCTTGGCGATGTACACAGCATGTGTGCTCAGGCCATGCGTGTATATGAGACCTT	375
OY	1020	CACCAAGCCCCAGTGGCAGGAGACTCCAGACGTACATCTCCAGTGGGAGAGCAGAGATGT	1079
Db	374	CACCAAGCCCCAGTGGCAGGAGACTCCAGACGTACATCTCCAGTGGGAGAGCAGAGATGT	315
OY	1080	CCAGGATTCAGATGCAGATNTTGTATGAGAACTTGTATCGAGTCCAAAGATGATTTAGTGAG	1139
Db	314	CCAGGATTCAGATGCAGATNTTGTATGAGAACTTGTATCGAGTCCAAAGATGATTTAGTGAG	255
OY	1140	CTGGGCTTNGCCAGGAGATCTGGGGACAAAGAAAGCAGATTTTCTGTATT	1188
Db	254	CTGGGCTTNGCCAGGAGATCTGGGGACAAAGAAAGCAGATTTTCTGTATT	206

RESULT 4

PCT-US01-04098A-2212/c
 : Sequence 2212, Application PC/TUS0104098A
 : GENERAL INFORMATION:
 : APPLICANT: Hyseq, Inc.
 : TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
 : FILE REFERENCE: 21272-029
 : CURRENT APPLICATION NUMBER: PCT/US01/04098A
 : CURRENT FILING DATE: 2001-02-05
 : PRIOR APPLICATION NUMBER: Not Yet Assigned

```

? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: 09/728,422
? PRIOR FILING DATE: 2000-11-30
? PRIOR APPLICATION NUMBER: 09/693,325
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 09/663,561
? PRIOR FILING DATE: 2000-09-15
? PRIOR APPLICATION NUMBER: 09/654,936
? PRIOR FILING DATE: 2000-09-01
? PRIOR APPLICATION NUMBER: 09/620,325
? PRIOR FILING DATE: 2000-07-19
? PRIOR APPLICATION NUMBER: 09/598,075
? PRIOR FILING DATE: 2000-06-20
? PRIOR APPLICATION NUMBER: 09/560,875
? PRIOR FILING DATE: 2000-04-27
? PRIOR APPLICATION NUMBER: 09/496,914
? PRIOR FILING DATE: 2000-02-03
? NUMBER OF SEQ ID NOS: 3960
? SOFTWARE: Custom
? SEQ ID NO 2212
? LENGTH: 1398
? TYPE: DNA
? ORGANISM: Homo sapiens
PCT-US01-04098A-2212
```

Query Match 97.6%; Score 1160; DB 1; Length 1398;
Best Local Similarity 99.4%; Pred. No. 2,6e-278;

Matches 1182; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

```

QY 1 GGGTTCTGGAGGAGCCGCTCCGCTGCTGTTGTTCCGGAAGTCTGCGGGGTGG 60
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 1393 GGGTTCTGGAGGAGCCGCTCCGCTGCTGTTGTTCCGGAAGTCTGCGGGGTGG 1334
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
QY 61 GAAATGCTGGCGGGGGGGGGGGGCACTGGGGGCCCTTTTCTAGAGGGCTCTACT 120
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 1333 GAAATGCTGGCGGGGGGGGGGGGCACTGGGGGCCCTTTTCTAGAGGGCTCTACT 1275
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
QY 121 GGGTTCTGGCGGGGGGGGGGGGCACTGGGGGCCCTTTTCTAGAGGGCTCTACT 179
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 1274 GGGTTCTGGCGGGGGGGGGGGGCACTGGGGGCCCTTTTCTAGAGGGCTCTACT 1215
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
QY 180 GTTTCGTCGCGCAGCAGAGAGGCGCTGGGGTGTGAGCGCAATGGGCCGATTCACCGGATCCT 239
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 1214 GTTTCGTCGCGCAGCAGAGAGGCGCTGGGGTGTGAGCGCAATGGGCCGATTCACCGGATCCT 1155
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
QY 240 GAGGCGCTGGTTGAACATCTCTATCCCTGTGTTAGACCCGATCCGATATGTGAGAGTCT 299
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 1154 GAGGCGCTGGTTGAACATCTCTATCCCTGTGTTAGACCCGATCCGATATGTGAGAGTCT 1095
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
QY 300 CAAGAAATTTGTCATCAACGTCGCTGAGCAGTGGCTGTGACTCTCGACAATGTAACTCT 359
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 1094 CAAGAAATTTGTCATCAACGTCGCTGAGCAGTGGCTGTGACTCTCGACAATGTAACTCT 1035
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
QY 360 GCAATTCGATGAGTCTCTTACTCTGGGCATCATGAGACCTTTACAAGGCAAGTACGGTGT 419
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 1034 GCAATTCGATGAGTCTCTTACTCTGGGCATCATGAGACCTTTACAAGGCAAGTACGGTGT 975
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
QY 420 GAGGAGCCCTGATATGCGGTCAACAGTGTCAAAACAACATGATGATCAAGTCTCGG 479
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 974 GAGGAGCCCTGATATGCGGTCAACAGTGTCAAAACAACATGATGATCAAGTCTCGG 915
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
QY 480 CAAACTCTCTGTTGGACAAGTCTTCCGGGAACGGAGTCCCTGAATGCCAGCATTTGTGA 539
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 914 CAAACTCTCTGTTGGACAAGTCTTCCGGGAACGGAGTCCCTGAATGCCAGCATTTGTGA 855
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
QY 540 TGGCATTCACCAAGTCTGCTGCTGCTGGGGTATCCGCTGCCCTTCGTTATGAGATCAAGGA 599
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 854 TGGCATTCACCAAGTCTGCTGCTGCTGGGGTATCCGCTGCCCTTCGTTATGAGATCAAGGA 795
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
QY 600 TATTCATGTGCCACCCCGGGTGAAGAAGTGTATGACAGATGACAGTGGAGGAGGCGG 659
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 794 TATTCATGTGCCACCCCGGGTGAAGAAGTGTATGACAGATGACAGTGGAGGAGGCGG 735
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
```

```

QY 660 GAAACGGGCCACAGTTCTAGAGTCTGAGGGAGCCGAGAGTCCGCCATCATGTGGCAGA 719
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 734 GAAACGGGCCACAGTTCTAGAGTCTGAGGGAGCCGAGAGTCCGCCATCATGTGGCAGA 675
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
QY 720 AGGAGAAACAGGCCACAGTCTCTGGCTCCGAGCAGAAAGGCTGTAACAGATAATCA 779
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 674 AGGAGAAACAGGCCACAGTCTCTGGCTCCGAGCAGAAAGGCTGTAACAGATAATCA 615
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
QY 780 GGGAGCAGAGAGGCCAGTCCAGTCTGGGGAAGGCCCAAGCTAAAGCTAAGCTATTTCG 839
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 614 GGGAGCAGAGAGGCCAGTCCAGTCTGGGGAAGGCCCAAGCTAAAGCTAAGCTATTTCG 555
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
QY 840 AATCTGGCTGACAGTCTTGACACACATTAATGAGATGACAGAGCTTCACTGACTGGCC 899
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 554 AATCTGGCTGACAGTCTTGACACACATTAATGAGATGACAGAGCTTCACTGACTGGCC 495
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
QY 900 CGAGCAGTATGTACAGCGGCTTTCCAAACCTGGCCCAAGGACTCCAAACATATCTACTGCC 959
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 494 CGAGCAGTATGTACAGCGGCTTTCCAAACCTGGCCCAAGGACTCCAAACATATCTACTGCC 435
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
QY 960 CTCCAACCCCTGGGATGTCAACAGCATGTTGGCTCAGGCGCATGGTGTATATGAGCCCT 1019
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 434 CTCCAACCCCTGGGATGTCAACAGCATGTTGGCTCAGGCGCATGGTGTATATGAGCCCT 375
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
QY 1020 CACCAAAAGCCCGAGTGCAGAGGACTCCAGACTCCTCCAGTGGAGCAGAGATGT 1079
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 374 CACCAAAAGCCCGAGTGCAGAGGACTCCAGACTCCTCCAGTGGAGCAGAGATGT 315
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
QY 1080 CAAAGGTACAGATGCAAGTTGATGAGAGACTTTGATGATGATCAAGATGATTAGTGAG 1139
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 314 CAAAGGTACAGATGCAAGTTGATGAGAGACTTTGATGATGATGATCAAGATGATTAGTGAG 255
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
QY 1140 CTGGGCTTNGCCAGAGTCTGGGGAACAAGCAAGATTTTCTGATT 1188
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
DB 254 CTGGGCTTNGCCAGAGTCTGGGGAACAAGCAAGATTTTCTGATT 206
    |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||

RESULT 5
US-09-471-275-865/C
? Sequence 865, Application US/09471275
? GENERAL INFORMATION:
? APPLICANT: Hyseq, Inc.
? TITLE OF INVENTION: Novel Contigs Obtained
? TITLE OF INVENTION: From Various Libraries
? FILE REFERENCE: 782
? CURRENT APPLICATION NUMBER: US/09/471,275
? EARLIER FILING DATE: 1999-12-23
? EARLIER APPLICATION NUMBER: US 09/235,076
? EARLIER FILING DATE: 1999-01-20
? EARLIER APPLICATION NUMBER: US 09/234,611
? EARLIER FILING DATE: 1999-01-22
? EARLIER APPLICATION NUMBER: US 09/240,371
? EARLIER FILING DATE: 1999-01-29
? EARLIER APPLICATION NUMBER: US 09/277,227
? EARLIER FILING DATE: 1999-03-25
? EARLIER APPLICATION NUMBER: US 09/271,490
? EARLIER FILING DATE: 1999-03-18
? EARLIER APPLICATION NUMBER: US 09/293,972
? EARLIER FILING DATE: 1999-04-15
? EARLIER APPLICATION NUMBER: US 09/274,861
? EARLIER FILING DATE: 1999-03-23
? EARLIER APPLICATION NUMBER: US 60/125,453
? EARLIER FILING DATE: 1999-03-19
? EARLIER APPLICATION NUMBER: US 60/126,605
? EARLIER FILING DATE: 1999-03-26
? EARLIER APPLICATION NUMBER: US 09/306,350
? EARLIER FILING DATE: 1999-05-07
? EARLIER APPLICATION NUMBER: US 09/399,720
? EARLIER FILING DATE: 1999-09-21
? EARLIER APPLICATION NUMBER: US 09/404,284
? EARLIER FILING DATE: 1999-09-21
? EARLIER APPLICATION NUMBER:
```

EARLIER FILING DATE: 1999-12-16
NUMBER OF SEQ ID NOS: 10451
SOFTWARE: PL_CT_genes Version 1.0
SEQ ID NO 865
LENGTH: 1398
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1296)..(264)
OTHER INFORMATION: similar to g12984585 in the genepept database release 114,
OTHER INFORMATION: Run with FASTX 3.3t00, default parameters
US-09-471-275-865

Query Match 97.6% Score 1160; DB 18; Length 1398;
Best Local Similarity 99.4%; Pred. No. 2.6e-278;
Matches 1182; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

1 GGCCTCTGGAGACNACCCCTCCGCTCCTCTGCTGTTGTTCCGAGAGTCCGCTCGCGCGGTG 60
1393 GGCCTCTGGAGACNACCCCTCCGCTCCTCTGCTGTTGTTCCGAGAGTCCGCTCGCGCGGTG 1334
61 GAAATGCTGGCG 120
1333 GAAATGCTGGCG 1275
121 GGCCTCTGGCG 179
1274 GGCCTCTGGCG 1215
180 GTTCTGCCCGCAGCAGAGAGCCCTGGGTGTGAGCGAATGGCCCGATTCACCGCATCT 239
1214 GTTCTGCCCGCAGCAGAGAGCCCTGGGTGTGAGCGAATGGCCCGATTCACCGCATCT 1155
240 GGAGCGCTGGTTGAACATCCTCATCCCTGTGTAGACCGGATCCGATGTGTCAGAGTCT 299
1154 GGAGCGCTGGTTGAACATCCTCATCCCTGTGTAGACCGGATCCGATGTGTCAGAGTCT 1095
300 CAAGAAATTTTCATCAACAGTCTGAGCAGTCCGCTGTGATCTTCGACAAATGTAACTCT 359
1094 CAAGAAATTTTCATCAACAGTCTGAGCAGTCCGCTGTGATCTTCGACAAATGTAACTCT 1035
360 GCAATTCATGATGATCTTACCTGCGCATCATGACACCTTACAAAGGCAAGCTACGCTGT 419
1034 GCAATTCATGATGATCTTACCTGCGCATCATGACACCTTACAAAGGCAAGCTACGCTGT 975
420 GGAGGACCCCTGATGATGCGGTACCCAGTACGTCAAACACCATGATGATGATGATGATG 479
974 GGAGGACCCCTGATGATGCGGTACCCAGTACGTCAAACACCATGATGATGATGATGATG 915
480 CAAACTCTCTNTGAGCAAGTCTTCCGGGAGCGGAGTCCCTGTAATGCCAGATTGTGGA 539
914 CAAACTCTCTNTGAGCAAGTCTTCCGGGAGCGGAGTCCCTGTAATGCCAGATTGTGGA 855
540 TGGCATCAACCAAGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 599
854 TGGCATCAACCAAGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 795
600 TATTCATTCGTCACCCCGGCTGAAAGTCTATGACAGATTCAGATGAGGAGGAGGAGGCG 659
794 TATTCATTCGTCACCCCGGCTGAAAGTCTATGACAGATTCAGATGAGGAGGAGGAGGCG 735
660 GAAACGGGCGACAGTCTAGAGTCTAGAGGAGCCGAGAGTCCGCCATCAATGTGCGAGA 719
734 GAAACGGGCGACAGTCTAGAGTCTAGAGGAGCCGAGAGTCCGCCATCAATGTGCGAGA 675
720 AGGGAAGAAACAGAGCCAGATCTTGGCCTTCGAGCAGAAAGGCTGAGACAGATAATCA 779
674 AGGGAAGAAACAGAGCCAGATCTTGGCCTTCGAGCAGAAAGGCTGAGACAGATAATCA 615
780 GCGAGCAGAGAGAGCCAGTCAATCTTGGGGAAGGCCAAGGCTAAAGCTCAACTATTTC 839

Db 614 GGCACAGAGAGAGGCCAGTGCAGTTCGCGCAGAGCCCAAGGCTAAAGCTGAAGCTATTGC 555
Qy 840 AATCTGCGCTCAGCTCTGACACACATATGAGATGACAGAGCTTCACTGATCTGTC 899
Db 554 AATCTGCGCTCAGCTCTGACACACATATGAGATGACAGAGCTTCACTGATCTGTC 495
Qy 900 CGAGCAGTATGACAGCGGCTCTCCAACTGGCCCAAGAGCTCCAACTATCTACTGTC 959
Db 494 CGAGCAGTATGACAGCGGCTCTCCAACTGGCCCAAGAGCTCCAACTATCTACTGTC 435
Qy 960 CTCCAACTGAGCAGTCTCACCAGCAGTGTGCTCAGCCATGAGGCTGATATGAGCCCT 1019
Db 434 CTCCAACTGAGCAGTCTCACCAGCAGTGTGCTCAGCCATGAGGCTGATATGAGCCCT 375
Qy 1020 CACCAAGCCCGCAGTCCAGGAGCTCCAGACTCACTCTCCAGTGGAGCAGCAGAGATGT 1079
Db 374 CACCAAGCCCGCAGTCCAGGAGCTCCAGACTCACTCTCCAGTGGAGCAGCAGAGATGT 315
Qy 1080 CCAGGCTACAGTGCAGCTTGTGATGAGGAACTGTGATGAGTCAAGATGATGATGAG 1139
Db 314 CCAGGCTACAGTGCAGCTTGTGATGAGGAACTGTGATGAGTCAAGATGATGATGAG 255
Qy 1140 CTGGCTTNGCCAGGAGTCTGGGAGCAAGAGCAGATTTTCTGATTT 1188
Db 254 CTGGCTTNGCCAGGAGTCTGGGAGCAAGAGCAGATTTTCTGATTT 206

RESULT 6
US-09-496-914A-8410/C
Sequence 8410, Application US/09496914A
GENERAL INFORMATION:
APPLICANT: Tang, Yuanhua T.
APPLICANT: Tillinghast, John
APPLICANT: Slinku, Ankura
APPLICANT: Liu, Chenghua
APPLICANT: Drmanec, Radoje T.
TITLE OF INVENTION: Novel Contigs Obtained
FILE REFERENCE: From Various Libraries
CURRENT FILING DATE: 2000-02-03
PRIOR APPLICATION NUMBER: US 09/353,690
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: US 09/034,341
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 09/045,400
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: US 09/321,214
PRIOR FILING DATE: 1999-05-26
PRIOR APPLICATION NUMBER: US 09/131,598
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: US 09/431,517
PRIOR FILING DATE: 1999-11-01
PRIOR APPLICATION NUMBER: US 09/328,351
PRIOR FILING DATE: 1999-06-04
PRIOR APPLICATION NUMBER: US 09/332,782
PRIOR FILING DATE: 1999-06-14
PRIOR APPLICATION NUMBER: US 09/235,076
PRIOR FILING DATE: 1999-01-20
PRIOR APPLICATION NUMBER: US 09/234,611
PRIOR FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: US 09/346,956
PRIOR FILING DATE: 1999-07-02
PRIOR APPLICATION NUMBER: US 09/362,510
PRIOR FILING DATE: 1999-07-27
PRIOR APPLICATION NUMBER: US 09/240,371
PRIOR FILING DATE: 1999-01-29
PRIOR APPLICATION NUMBER: US 09/248,797
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 09/271,490
PRIOR FILING DATE: 1999-03-18
PRIOR APPLICATION NUMBER: US 09/293,972
PRIOR FILING DATE: 1999-04-15

Matches 1182: Conservative 0; Mismatches 5; Indels 2; Gaps 2;

```
OY 1 GGCCTCTGGAGACNACCGCTCCGCTGCTCGTTGGTCCGAGAGCGCGTGGCGGCTGG 60
DB 1393 GGCCTCTGGAGACGACCGCTCCGCTGCTCGTTGGTCCGAGAGCGCGTGGCGGCTGG 1334
OY 61 GAAATGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
DB 1333 GAAATGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1275
OY 121 GGCCTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 179
DB 1274 GGCCTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1215
OY 180 GTTCGTGGCGCGAGAGAGAGCGCGTGGTGGAGCGAATGGCGCGATTCACCGATCC 239
DB 1214 GTTCGTGGCGCGAGAGAGAGCGCGTGGTGGAGCGAATGGCGCGATTCACCGATCC 1155
OY 240 GGAGCGCTGGTTGAACATCCTCATCCCTGTGTAGACCGGATCCGTTATGTGAGAGTCT 299
DB 1154 GGAGCGCTGGTTGAACATCCTCATCCCTGTGTAGACCGGATCCGTTATGTGAGAGTCT 1095
OY 300 CAGGAAATGTGATCATCAAGCTGCTGAGAGAGTGGCTGTGACTCTGGACAATGTAACTCT 359
DB 1094 CAGGAAATGTGATCATCAAGCTGCTGAGAGAGTGGCTGTGACTCTGGACAATGTAACTCT 1035
OY 360 GCAAAATCGATGAGAGTCTTACCTGGCATCATGAGACCGTTACAGCGCAAGCTAGCGTGT 419
DB 1034 GCAAAATCGATGAGAGTCTTACCTGGCATCATGAGACCGTTACAGCGCAAGCTAGCGTGT 975
OY 420 GGAGGACCTGTGATGCGCGTCAACCGAGTACCTCAACACCATGAGATCAGAGCTCGG 479
DB 974 GGAGGACCTGTGATGCGCGTCAACCGAGTACCTCAACACCATGAGATCAGAGCTCGG 915
OY 480 CAAACTCTCTNTGACAAAGTCTTCCGGGAGAGGAGTCCCTGAGAGCCAGCATTTGGA 539
DB 914 CAAACTCTCTNTGACAAAGTCTTCCGGGAGAGGAGTCCCTGAGAGCCAGCATTTGGA 855
OY 540 TGCATCAACAAGCTGCTGAGTGTGGGGTATCCGCTGCTTCGTTATGAGATCAAGA 599
DB 854 TGCATCAACAAGCTGCTGAGTGTGGGGTATCCGCTGCTTCGTTATGAGATCAAGA 795
OY 600 TATCATGTGCCACCCCGGGTGAAGAGTCTATGACAGATGAGAGTGGAGAGCGCGG 659
DB 794 TATCATGTGCCACCCCGGGTGAAGAGTCTATGACAGATGAGAGTGGAGAGCGCGG 735
OY 660 GAAAGGGGCGACAGTCTGAGAGTGTGAGGGGACCGGAGAGTGGCGATCAATGTGACGA 719
DB 734 GAAAGGGGCGACAGTCTGAGAGTGTGAGGGGACCGGAGAGTGGCGATCAATGTGACGA 675
OY 720 AGGGAAGAAACAGGCCAGATCCTGGCTCCGAGAGAGAGAGAGAGAGAGAGAGAGTAATCA 779
DB 674 AGGGAAGAAACAGGCCAGATCCTGGCTCCGAGAGAGAGAGAGAGAGAGAGAGTAATCA 615
OY 780 GGCAGAGAGAGAGGCCAGTGTGAGTGTGGCGAAGGCCAAGGCTAAAGCTGAAGCTATTTCG 839
DB 614 GGCAGAGAGAGAGGCCAGTGTGAGTGTGGCGAAGGCCAAGGCTAAAGCTGAAGCTATTTCG 555
OY 840 AATTCCTGGTGGAGCTGTACACAAATATATGAGATGAGAGAGTGTCACTGACTGTGGC 899
DB 554 AATTCCTGGTGGAGCTGTACACAAATATATGAGATGAGAGAGTGTCACTGACTGTGGC 495
OY 900 CGAGCAGTATGTACAGCGCTTCTCCAACTGGCGAAGAGCTCCAAAGCTATCTACTGGC 959
DB 494 CGAGCAGTATGTACAGCGCTTCTCCAACTGGCGAAGAGCTCCAAAGCTATCTACTGGC 435
OY 960 CTCGCAACCTTGGCGATGTACACAGATGTGTGCTCAGGCGCATGGAGTGTATGTAGAGCCCT 1019
DB 434 CTCGCAACCTTGGCGATGTACACAGATGTGTGCTCAGGCGCATGGAGTGTATGTAGAGCCCT 375
OY 1020 CACCAAAAGCCCAAGTGTCCAGGAGATCACAATCTCTCACTGAGGAGCAGCAGAGATGT 1079
DB 374 CACCAAAAGCCCAAGTGTCCAGGAGATCACAATCTCTCACTGAGGAGCAGCAGAGATGT 315
```

```
OY 1080 CCAGGCTACAGATGCAAGTNTTGATGAGAGACTGTGATCGAGTCAAGATGATTAGTGAG 1139
DB 314 CCAGGCTACAGATGCAAGTNTTGATGAGAGACTGTGATCGAGTCAAGATGATTAGTGAG 255
OY 1140 CTGGGCTTGGCCAGGAGGAGTCTGTGGGACACAGAAGACGAGATTTCTGATT 1188
DB 254 CTGGGCTTGGCCAGGAGGAGTCTGTGGGACACAGAAGACGAGATTTCTGATT 206
```

RESULT 8

```
US-09-338-467-941
; Sequence 941, Application US/09338467
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; FILE REFERENCE: MLN98-25P
; CURRENT APPLICATION NUMBER: US/09/338,467
; EARLIER FILING DATE: 1999-06-22
; EARLIER APPLICATION NUMBER: 60/090,177
; NUMBER OF SEQ ID NOS: 1022
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 941
; LENGTH: 1384
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-338-467-941
```

Query Match 97.2%; Score 1155; DB 17; Length 1384;

Best Local Similarity 99.4%; Pred No. 4.5e-277;

Matches 1177: Conservative 0; Mismatches 5; Indels 2; Gaps 2;

```
OY 6 CTGGAGACNACCGCTCCGCTGCTCGTTGGTCCGAGAGTGGCTGGCGGAGAAAT 65
DB 4 CTGGAGAGCGAGCGCTCGCTCGTCTGCTCGTGGAGTGGCTGGCGGAGAAAT 63
OY 66 GCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 125
DB 64 GCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 122
OY 126 CTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 184
DB 123 CTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 182
OY 185 TGCAGGAGAGAGCGCTGGTGGTGGAGCGAATGGGCGGATTCACCGGATCTGGAGC 244
DB 183 TGCAGGAGAGAGCGCTGGTGGTGGAGCGAATGGGCGGATTCACCGGATCTGGAGC 242
OY 245 CTGGTTGAAACATCCATCCCTGTGTGATGACCGGATTCGATATGTGAGATCTCAAG 304
DB 243 CTGGTTGAAACATCCATCCCTGTGTGATGACCGGATTCGATATGTGAGATCTCAAG 302
OY 305 AATTTGATCAACAGTGGCTGAGAGAGTGGCTGTGACTCTGACAAATGTAACTCTCAAA 364
DB 303 AATTTGATCAACAGTGGCTGAGAGAGTGGCTGTGACTCTGACAAATGTAACTCTCAAA 362
OY 365 TCGATGAGTCTCTTAACTGAGCATATGAGACCTTAAAGGCAAGCTACGTTGGAGG 424
DB 363 TCGATGAGTCTCTTAACTGAGCATATGAGACCTTAAAGGCAAGCTACGTTGGAGG 422
OY 425 ACCCTGAGTATGCTGACACCGAGCTAGCTCAAAACAACATGAGATGAGGCTGGCAAC 484
DB 423 ACCCTGAGTATGCTGACACCGAGCTAGCTCAAAACAACATGAGATGAGGCTGGCAAC 482
OY 485 TCTCTTGGACAAGCTCTCCGGGAACGGGAGTCCCTGAATGCGACGATGTGGATGCCA 544
DB 483 TCTCTTGGACAAGCTCTCCGGGAACGGGAGTCCCTGAATGCGACGATGTGGATGCCA 542
OY 545 TCACCAACGCTGACTGCTGGGATATCCGCTGCTTCGTTATGAGATCAAGATATCC 604
```



```

Db 1083 gtaccagatgcaagctctgtatggaactgacgagtcacaagatgagttgagagctgg 1142
Oy 1145 CTTCGACAGAGTCTGGGAGCAGAGAACAGATTTTCTGATT 1188
Db 1143 ctggcaggagagctctgggacaaagaaagagatttctcgtatc 1186

RESULT 10
US-09-649-162-7309
; Sequence 7309, Application US/09649162
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; APPLICANT: Richardson, Jennifer
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600,1181-001
; CURRENT APPLICATION NUMBER: US/09/649,162
; CURRENT FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: 60/151,057
; PRIOR FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 9990
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 7309
; LENGTH: 1384
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-649-162-7309

Query Match 97.2%; Score 1155; DB 25; Length 1384;
Best Local Similarity 99.4%; Pred. No. 4,5e-277;
Matches 1177; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

Oy 6 CTGGAGACGACCGCTCCGTCGTCGTTGTCGCGAGAGTGCCTGCGGCGGTGGGAAT 65
Db 4 ctggagagcagccgctcgcgtctcgtgttcggaagtcgctcgcgcggtgggaaat 63
Oy 66 GCTGGCGCGCGCGCGGCGGCGACACTGGGGCCCTTTGCTGAGGGGCTCTACTAGGCTT 125
Db 64 gctggcgcgcgcgcgcgcg-ggggactggggccctttgtcgggggctctcactggtc 122
Oy 126 CTGGCGCGGCTCG-GCCGCGCTCTCTGTGATTGCCCGGAACCGTGGTACTGTTGC 184
Db 123 ctggcgcgcgctcgcgcgcgcctcctcgtgattgccccgaaacacggtgactgctg 182
Oy 185 TGCGGACAGAGGAGCGGTGGTGGAGCGAATGGCCGATTCACCGGATCTCGAGAC 244
Db 183 tggcgagcagaagagcgctgggtggtgagcgaaatgggacgattccacggtatcggagc 242
Oy 245 CTGCTTTGAACATCCTCATCCCTGTGTGAGACCGGATCCGATATGTCAGAGTCTCAAG 304
Db 243 ctggtttgaaatcctcctcctgtgtgtagaccgataccgataatgtagagagtcacaag 302
Oy 305 AATTGTATCACTGCTGCTGACAGTGGCTGTGACTGTGACAAATGTAATCTTGCAAA 364
Db 303 aattgtatcatacagctgctcggagcagctggtgactctcacaatgtaactctgcaaa 362
Oy 365 TCGATGAGATCCTTTACCTGACGCGCATATGAGCCTTACAAAGCAGAGTCCGGTGGAG 424
Db 363 tggatggagctccttaactctgagcacaatgagaccttaagaagcctcggtggtgag 422
Oy 425 ACCCTGAGTATCGCTACACCAAGCTAGCTCAAAACAACATGAGATCAGAGCTCGGAAC 484
Db 423 accctgagatgctcgctacaccagctagctcaaaacacatgagatcagagctcgggaaac 482
Oy 485 TCTCTTTGACAAAGTCTTCCGGGAACGGAGTCCCTGATATCCAGACATTTGATGATCCA 544
Db 483 tctctcttggacaagaatctctcgggaagagagctccctgagatgcccagcatgtgagatcca 542
Oy 545 TCAACCAAGCTCTGACTGCTGGGGATCCGCTGCTGCTGATGATGATCAAGGATATCC 604
Db 1143 ctggcaggagagctctgggacaaagaaagagatttctcgtatc 1186
```

```

Db 543 tcaaccaagctgctgactgctggtggtatccgctcctccgttatgagatcaagaatattcc 602
Oy 605 ATGTGACACCCCGGGGTAAGAAGTCTATGCAGATCAGATGGAGGACAGCGCGGAAC 664
Db 603 atgtgacaccccggttgaaagatctatgcaatgcaatgagtgagagcgagcggtgaaac 662
Oy 665 GGGCCACAGTTTCTAGAGTGTGAGGGACCCGAGAGTCCGCCATCAATGTGCGAAGGA 724
Db 663 gggccacagctctagatgctgagggagaccgagagctcgccatcaatgtgagagaagg 722
Oy 725 AGAAGACGGCCCAAGATCCTGCGCCTCCGAAGCAGAAAGGCTCAACAGATTAATCAGCG 784
Db 723 agaaacagccccagatctccgctccgaagcagaagaagctgaacagataaaatcagcg 782
Oy 785 CAGAGAGGSCAGTGCAGTGTGCGGAAGGCCAAGCTTAACCTGATGATTCGATCC 844
Db 783 cagagagagccagctcagctcgtcggaagccaaagctaaagctgaatctcgaatcc 842
Oy 845 TGGCTGCAGCTCTGACACAACTAATGAGATGCAAGCAGCTTCACTGACTGTGGCGAGC 904
Db 843 tggctgcagctctgacacaataatgagatgcaagcagcttcactgactgtggtcgagc 902
Oy 905 ACTATGTCAAGCCGCTTCTGCAAACTGGCCCAAGAGACTCCAAACTATCTACTGCCCC 964
Db 903 agtatgtcagcggtctcctcaactggtccaaagactccaacatctcctcctcca 962
Oy 965 ACCCTGCGCATGTCAACAGCATGATGTGAGTCCAGGCCATGAGTATATGAGCGCTCACCA 1024
Db 963 accctgagatgctcaacagcagctggtggtccagagccatggtatagagccctccaca 1022
Oy 1025 AAGCCCACTGCCAGGAGACTCCAGACTCACTCTCCAGTGGGAGCAGAGATGTCCAG 1084
Db 1023 aagcccaactgccaagagactccagactcctcctcagtggtgagcagcagaatctcag 1082
Oy 1085 GTACAGATCAGATNTGATGAGAACTGATGAGTCAAGTGAATGATGTGAGCTGGG 1144
Db 1083 gtacagatgcaagctctgtgagagactgacatgcaatgagatgagttgagagctgg 1142
Oy 1145 CTTCGACAGAGTCTGGGAGCAGAGAACAGATTTTCTGATT 1188
Db 1143 ctggcaggagagctctgggacaaagaaagagatttctcgtatc 1186

RESULT 11
US-09-652-109-7925
; Sequence 7925, Application US/09652109
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600,1180-001
; CURRENT APPLICATION NUMBER: US/09/652,109
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,128
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 10105
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 7925
; LENGTH: 1384
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-109-7925

Query Match 97.2%; Score 1155; DB 25; Length 1384;
Best Local Similarity 99.4%; Pred. No. 4,5e-277;
Matches 1177; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

Oy 6 CTGGAGACGACCGCTCCGTCGTCGTTGTCGCGAGAGTGCCTGCGGCGGTGGGAAT 65
Db 4 ctggagagcagccgctcgcgtctcgtgttcggaagtcgctcgcgcggtgggaaat 63
Oy 66 GCTGGCGCGCGCGGCGGCGGCGACACTGGGGCCCTTTTGTGAGGGGCTCTACTAGGCTT 125
```


Db	663	gaggccacagttctcagagctctcaggggagaccgagatctcgccatcaatctgtgcagaaaggga	722
Qy	725	AGAAACAGCCCCAGATCCCGGGCTCCGAAAGCGAAAGAGCTGAACAGATTAATCAGCAG	784
Db	723	agaaabaagggcccgagatccctgagctcccgaaagccgaaagaagctcgaaacagataaataagcag	782
Qy	785	CAGAGAGGCCAATCGATGTTCTGGCGAAGCGCAAGCTTAAAGCTTGAAGCTATTGGAATCC	844
Db	783	cagagagagagccagctgagatcttcgycgaaagggccaagggctaaagctgaaagctatcgaatcc	842
Qy	845	TGGCTCAGCTCTGCACACAACATATGAGATGCAAGAGCTTCACTGACTGTGGCCGACG	904
Db	843	tggctcgagagctcgacaacaataatgagatgacagagctctccatctgactgycgagc	902
Qy	905	AGATATGACAGCGGTTCTCCAAATCTGGCCCAAGAGACTTCCAAACACTTCTACTGCTCTCA	964
Db	903	agatatgacagcgagctctccaaactctgcccgaagatctcaaacctactctactgcccctca	962
Qy	965	ACCGTGGCGATGTCACACAGCATGTTGGCTTCAGGCCATGGGTGTATATGAGACCTCAGCA	1024
Db	963	accctggcgatgtccacaagaatctgycctcgaagccaatctgatatagaaacccctcacca	1022
Qy	1025	AAGCCCAATGCGCAGGAGATCCGAGACTTCATCTCCAGTGGGAGCGAGAGATGTCAGG	1084
Db	1023	aagccccaatgcccagaggaaccccgagactccatcccgctggagggcagcgagagatgtccag	1082
Qy	1085	GTACAGATGCAAGTNTTGGATGAGAGAACTTGATTCGATCGAATCAAGATGAGTTAGTGGAGCTGGG	1144
Db	1083	gtacaagtgtcaagatctctgactgagaggaactctgactcgatcgaagtgaagtctagctgagctgg	1142
Qy	1145	CTTTCGACAGGAGTCTGGGGACAAGCAAGCAAGATTTTCTGTATT	1188
Db	1143	ctctggccagagagatctctggggaacaagaaacagatcttctctgatt	1186

```

RESULT 13
US-09-652-816-7419
; Sequence 7419, Application US/09652816
; GENERAL INFORMATION:
; APPLICANT: Gutierrez-Ramos, Jose-Carlos
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1177-001
; CURRENT APPLICATION NUMBER: US/09/652,816
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,111
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 9647
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7419
; LENGTH: 1384
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-816-7419

```

Query Match	97.28;	Score 1155;	DB 25;	Length 1384;
Best Local Similarity	99.48;	Pred. No. 4.5e-277;		
Matches 1177;	Conservative 0;	Mismatches 5;	Indels 2;	Gaps 2

QY 6 CTGGGAGCAGCAGCGCTCCGCTGTCTCCGTTGGTTCCGAGAGTCGCTGGCGCGTGGCGAAAT 65
 QY 5 CTGGGAGCAGCAGCGCTCCGCTGTCTCCGTTGGTTCCGAGAGTCGCTGGCGCGTGGCGAAAT 65
 Db 4 CTGGGAGCAGCAGCGCTCCGCTGTCTCCGTTGGTTCCGAGAGTCGCTGGCGCGTGGCGAAAT 63
 QY 66 GCTGGCGCGCGCGCGCGCGGACACTGGGCGCCCTTTTGGTGAAGGCGCTCTCACTGGCTT 125
 Db 64 GCTGGCGCGCGCGCGCGGACACTGGGCGCCCTTTTGGTGAAGGCGCTCTCACTGGCTT 122
 QY 126 CTGGCGCGCGCTCCG-CGCGCGCTCTCTTGATTGGACCCGAAACACCGTGTACTGTTG 184
 Db 123 CTGGCGCGCGCTCCGCGCGCGCGCTCTTGATTGGACCCGAAACACCGTGTACTGTTG 182

Qy	185	TCGCCGACGAGGAGCGCTGGGTGGTGAGAGCAATATGGCCGCAATTCCACCGGAGTCCCTGGAGC	244
Db	183	tgccgcacaggaagagccctgggctggctggagcgaaatgggcccgaattccacggagctccggagc	242
Qy	245	CTGGTTTGAACATCTCTCAATCCCTGTGTGTTAGACGGGATCCGATATGTGCACTTCTCAAGC	304
Db	243	ctggtttgaacactctctcaatccctgtgttagaccgagatccgatatgtgcagagcttcaagc	302
Qy	305	AAATTGTCAATCAAGCTGCCTGAGCAAGTCGGCTGTGACTTCGACAAATGTAACTCTGGCAA	364
Db	303	aaattgtcatcaacgctgcctggagcagctcgycctgtgactctgcgacaatgtlaactctggcaaa	362
Qy	365	TGCGATGAGATCCTTTAACTGGCCATCATGAGACCCCTTAAAGGCAAGCTACGGTGTGGAGC	424
Db	363	tgcgatgagatccttttaactggccatcatgagacccctttaaaggcaagctacggtgtggagc	422
Qy	425	ACCCCTGAGTATCCCGTCAACCCAGCTTAAGCTTAACAAACCAATGAGATCAGAGCTCGGCAAC	484
Db	423	acccctgagatgctccgtcacaccagctcagctcaaaacacatgataagctctggcaaac	482
Qy	485	TCCTCTTGACAAAGTCTTCGGGGAAGGGAGTCCCTGATATGGCAGCACTTGTGGATGCCA	544
Db	483	tctctctgacaagaagctcttcgggaacgggagctccctgaaatgcagcatgtgtgattgccaa	542
Qy	545	TCACCAACAGCTGCTGACTGCTGGGGTATCCGCTGCCTNGTTATGATGATCAAGATATCC	604
Db	543	tcaaccaagctctgactgcttggggtatccgcctgcgtctatgtagatcaagaatatcc	602
Qy	605	ATGTGCACACCCCGGATGAAAGATCTATGCAGATGAGAGTGAAGGAGAGCGGCGGAAAC	664
Db	603	atgtgcacaccccggtgaaagatctatgcagatgtagatgagagcgagcgycggaaac	662
Qy	665	GGGCCACAGTTTACAGTCTGAGGGGACCAGAAATGTCGGCATCATGTCGGCAGGAGGGA	724
Db	663	ggggccacagtttactagatctgaaagggaacccgaagctcgacataatgtggcagaaggga	722
Qy	725	AGAAACAGGCCCAGATCTGTGGCTCCGAAGCAGAAAAAGCTGAAACGATTAATCAGGAC	784
Db	723	agaaaaaggcccagatctgtgctccgaagcagaaaaagctgaaacagataaatacagcgag	782
Qy	785	CAGGAGAGGCCAGTGCAGTTCTTGCCGCAAGGCCAAAGCTAAAGCTGAAGCTAATTCGAAATCC	844
Db	783	caggagagagccagctgcagtctctggcgaaggccaaagctaaagctgaagctatctgaatcc	842
Qy	845	TGGCTGCAGCTTGACACACATTAATGGAATCAGAGCTTCACTGACTGTGGCCAGC	904
Db	843	tggctgcagcttgacacacaataatggaatcagagcttcactgactgtggccagc	902
Qy	905	AGTATGTCAAGCGCGTCTTCCAAACTGGCCCAAGCAACTCAACACTATCTACTGTGCCCTCCA	964
Db	903	agtatgtcagcgcttctccaacttggccaagagatccaacaactatctactgtgcccca	962
Qy	965	ACCTGCGCATGTACACAGCATGGTGGCTCAGGCGATGGGTGTATATGAGCCCTTCAACA	1024
Db	963	acctgcgcatgtacacagcatgggtggctcaggcgatgggtgtatatgagcccttcaaca	1022
Qy	1025	AAGCCCCAGTGCCAGGAGCTCCACACTCTTCCAGTGGGAGCAGACAGAGATGTCCAGG	1084
Db	1023	aagccccagtgccaggagctccacagctcactctcagttgggaagcagagaaatgtccagg	1082
Qy	1085	GTCACAGATGCAAGTNTTGATGAGGAACTTGAATCAGTCAAGTCAAGATCAGTTAGTGGACCTGGC	1144
Db	1083	gtcacagatgcagaagcttctgaaatggaactgtgacgagctcaagaatgagttagtggagctggg	1142
Qy	1145	CTTGGCCAGGAGAGTCTGGGGACAAGGAGACGATTTTCTGATT	1188
Db	1143	cttggccaggagagctctggggacaaggagacgagatcttccctcatt	1186

RESULT 14
US-09-652-913-8809
; sequence 8809, Application US/09652913
; GENERAL INFORMATION:

```

; APPLICANT: Falb, Dean R.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1174-001
; CURRENT APPLICATION NUMBER: US/09/652,913
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,107
; NUMBER OF SEQ ID NOS: 10833
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 8809
; LENGTH: 1384
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-913-8809
```

```

Query Match          97.2%  Score 1155; DB 25; Length 1384;
Best Local Similarity 99.4%  Pred. No. 4,5e-277;
Matches 1177; Conservative 0; Mismatches 5; Indels 2; Gaps 2;
```

```

QY 6 CTGGGAGACCGCTCCGTCGTCGTTGGTTCCGAGAGTGGCTGGCGCGTGGGAAT 65
DB 4 CTGGGAGCGCGCTCGCTCGCTCGTGTGTCGAGAGTCGTCGCGTGGGAAAT 63
QY 66 GCTGGCGCGCGCGCGCGCGGACACTGAGGCGCTTTGCTGAGGCGCTCTACTGCTT 125
DB 64 GCTGGCGCGCGCGCGCGCGGACACTGAGGCGCTTTGCTGAGGCGCTCTACTGCTT 122
QY 126 CTGGCGCGCTCCG-6CGCGCGCTCTCTGATTGGCCGAAACACCGTGTACTGTCG 184
DB 123 CTGGCGCGCTCCGCGCGCGCTCTCTGATTGGCCGAAACACCGTGTACTGTCG 182
QY 185 TGGCGGAGAGAGCGCGCTGGTGTGAGAGCAATGGCGCGGATTCACCGGATCTTGAGAC 244
DB 183 TGCGGAGAGAGAGCGCGCTGGTGTGAGAGCAATGGCGCGGATTCACCGGATCTTGAGAC 242
QY 245 CTGGTTTGAACATCCTCATCCTGTTGTAGACCGATCCGATATGACAGTCTCAAG 304
DB 243 CTGGTTTGAACATCCTCATCCTGTTGTAGACCGATCCGATATGACAGTCTCAAG 302
QY 305 AAATTGTATCAACAGTGCCTGAGACAGTGGCTGTGACTCTGACAAATGTACTTGTCAAA 364
DB 303 AAATTGTATCAACAGTGCCTGAGACAGTGGCTGTGACTCTGACAAATGTACTTGTCAAA 362
QY 365 TGGATGAGATCCTTTACCTGGCGCATCATGACCTTACAGCAACACTACGCTGTGAGAG 424
DB 363 TGGATGAGATCCTTTACCTGGCGCATCATGACCTTACAGCAACACTACGCTGTGAGAG 422
QY 425 ACCCTGAGTATGCGCTGACCCAGCTAGTCAAAACACATGAGATGAGAGTCTGGCAAC 484
DB 423 ACCCTGAGTATGCGCTGACCCAGCTAGTCAAAACACATGAGATGAGAGTCTGGCAAC 482
QY 485 TCTCTTGTGACAAAGTCTTCCGGGAAAGGAGTCCCTGTAATGCGACATTTGTGATGCCA 544
DB 483 TCTCTTGTGACAAAGTCTTCCGGGAAAGGAGTCCCTGTAATGCGACATTTGTGATGCCA 542
QY 545 TCAACCAACCTGCTGACCTGCGGTATGCCGCTGCTTCGTTATGATGATCAAGATATCC 604
DB 543 TCAACCAACCTGCTGACCTGCGGTATGCCGCTGCTTCGTTATGATGATCAAGATATCC 602
QY 605 ATGTGCGCACCGCGGTGAGAGAGTCTATGACAGTGCAGTGGAGGAGCGCGGAAAC 664
DB 603 ATGTGCGCACCGCGGTGAGAGAGTCTATGACAGTGCAGTGGAGGAGCGCGGAAAC 662
QY 665 GGGCCACACTTCTAGAGTGTAGGAGGAGCCGAGAGTGGCCATCAATGTGCGAGAGAGGA 724
DB 663 GGGCCACACTTCTAGAGTGTAGGAGGAGCCGAGAGTGGCCATCAATGTGCGAGAGAGGA 722
QY 725 AGAACAAGCGCCAGATCTCGGCTCCGAGCAAGAAAGGCTGAACGATTAATCAGGCGAG 784
DB 723 AGAACAAGCGCCAGATCTCGGCTCCGAGCAAGAAAGGCTGAACGATTAATCAGGCGAG 782
```

```

QY 785 CAGGAGAGCGCAGTGCAGTCTTGCGGAGGCCAAGGCTAAAGCTAATTCGAATCC 844
DB 783 CAGGAGAGCGCAGTGCAGTCTTGCGGAGGCCAAGGCTAAAGCTAATTCGAATCC 842
QY 845 TGGCTCAGCTCTGACACAAATATGAGATGACACACTTCTACTGAGTGGCGGAC 904
DB 843 TGGCTCAGCTCTGACACAAATATGAGATGACACACTTCTACTGAGTGGCGGAC 902
QY 905 AGTATGTCAGCGGCTTCTCCAAACCTGCGCAAGAGTCCAAACTATCTACAGCCCTGCA 964
DB 903 AGTATGTCAGCGGCTTCTCCAAACCTGCGCAAGAGTCCAAACTATCTACAGCCCTGCA 962
QY 965 ACCCTGCGATGTCACACAGATGTGCTCAGGCGCATGGGTATATGAGAGCCCTGACA 1024
DB 963 ACCCTGCGATGTCACACAGATGTGCTCAGGCGCATGGGTATATGAGAGCCCTGACA 1022
QY 1025 AAGCCCACTGCCAGGAGCTCCAGACTCACTCTCCAGTGGAGACAGACAGATGTCCAG 1084
DB 1023 AAGCCCACTGCCAGGAGCTCCAGACTCACTCTCCAGTGGAGACAGACAGATGTCCAG 1082
QY 1085 GTAACAGATGCAAGTNTGATGAGAGCACTGATCGAATCAAGATGATGAGAGCTGGG 1144
DB 1083 GTAACAGATGCAAGTNTGATGAGAGCACTGATCGAATCAAGATGATGAGAGCTGGG 1142
QY 1145 CTTCGCGAGGAGTCTGGGAGACAGAAAGCAATTTCTGATT 1188
DB 1143 CTTCGCGAGGAGTCTGGGAGACAGAAAGCAATTTCTGATT 1186
```

```

RESULT 15
US-09-652-914-7997
; Sequence 7997, Application US/09652914
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1193-001
; CURRENT APPLICATION NUMBER: US/09/652,914
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,112
; NUMBER OF SEQ ID NOS: 9677
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 7997
; LENGTH: 1384
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-914-7997
```

```

Query Match          97.2%  Score 1155; DB 25; Length 1384;
Best Local Similarity 99.4%  Pred. No. 4,5e-277;
Matches 1177; Conservative 0; Mismatches 5; Indels 2; Gaps 2;
```

```

QY 6 CTGGGAGACCGCTCCGTCGTCGTTGGTTCCGAGAGTGGCTGGCGCGTGGGAAT 65
DB 4 CTGGGAGCGCGCTCGCTCGCTCGTGTGTCGAGAGTCGTCGCGTGGGAAAT 63
QY 66 GCTGGCGCGCGCGCGCGGACACTGAGGCGCTTTGCTGAGGCGCTCTACTGCTT 125
DB 64 GCTGGCGCGCGCGCGCGGACACTGAGGCGCTTTGCTGAGGCGCTCTACTGCTT 122
QY 126 CTGGCGCGCTCCG-6CGCGCGCTCTCTGATTGGCCGAAACACCGTGTACTGTCG 184
DB 123 CTGGCGCGCTCCGCGCGCGCTCTCTGATTGGCCGAAACACCGTGTACTGTCG 182
QY 185 TGGCGGAGAGAGCGCGTGGTGTGAGAGCAATGGCGGATTCACCGGATCTTGAGAC 244
DB 183 TGGCGGAGAGAGCGCGTGGTGTGAGAGCAATGGCGGATTCACCGGATCTTGAGAC 242
QY 245 CTGGTTTGAACATCCTCATCCTGTTGTAGACCGATCCGATATGTCGAGATCTCAAG 304
DB 243 CTGGTTTGAACATCCTCATCCTGTTGTAGACCGATCCGATATGTCGAGATCTCAAG 302
```

OY 305 AAATTGTCATCAACCTGGCTGAGCAGTGGCTGTGACTCTCGACAATGTAACTCTGCAAA 364
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 303 Aaattlgcatcaaaagtcgctcgtgagcagtcgctgagctctcgaaatgtlaactctgcaaa 362
 OY 365 TCGATGGAGTCTCTTACCTGGGATCATGAGACCTTACAGGCAAGCTTACGGTGTGAGAG 424
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 363 tcgaatggagtccttacccttaccctcgacatcaatggacccttaccagctacagtggtgag 422
 OY 425 ACCCTGAGTATGCGCTGACCCAGTACGTCAAACCAACCATGATCAGAGCTCGGCAAC 484
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 423 accctgagtaatgcgtccacccagctagctcaacaacacatgagatcagagctcgcaaac 482
 OY 485 TCTCTNTGGCAAAAGCTTTCGGGAGACGGAGTCCCTGATGCCAGCATTTGTGATGCCA 544
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 483 tctctctgacaagaatctctccggaacggagatccctgaaatgcagcatltggaatgcca 542
 OY 545 TCAACCAAGCTGTGACGCTGCGGATATCCGCTGCTNCGTATGAGATCAAGGATATCC 604
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 543 tcaacaaagctgacgacgctgagatcccgctcccgctcctcgatcaatgagatcaagatattcc 602
 OY 605 ATGTGCCACCCCGGGTGAAGAAGTATGACAGTACAGTGCAGTGAGGACAGACGCGGAAAC 664
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 603 atgtgtccaccccggttgaaagatctatgacagatgcaggtgagagcagagcgcgaaac 662
 OY 665 GGGCCACAGTTCTTAGAGTCTGAGGGACCCGAGAGTCCGCCATCAATGTGCGCAGAGGGA 724
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 663 gggccacagctctcagaatctgagggaccggagatcgccatcaatgtgcagaaagga 722
 OY 725 AGAAMCAGGCCAGATTCCTGGGCTCCGAAAGCAAGAAAGGCTGAACAGATTAATCAGGCGAG 784
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 723 agaaacaaagcccaagatccctggcctccgagcagaagaagctgacacagataaaatcagcgag 782
 OY 785 CAGGAGAGGCCAGTGCAGTTCGTGGCGAAGGCCAAGGCTAAAGCTTAAGCTATTTCGAATCC 844
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 783 cagggagagggccagctcagctctcgggaagggccaaggtctaagctatcgaatcc 842
 OY 845 TGGCTGACAGCTCTGACACACATAATGAGATGACAGACTTCACTGACTGTGGCCGAGC 904
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 843 tggctgacagctctgacacacataaagatgagactcactgactgtgagcgagc 902
 OY 905 AGTATGTCAGGCGGCTTCCAAACCTGSCCAAGGACTCCAAACACTATTCCTACTGSCCTCCA 964
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 903 agtatgtcagcgctctccaaactggccaagactccaaactatccctacatcgccctcca 962
 OY 965 ACCCTGGCGATGTACACAGCATGTGCTCAGGCCATGGGTATATGAGGCCCTCACCA 1024
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 963 accctggagatgtacacagcatgtgtcagggccaatggtatataatgagccctcacca 1022
 OY 1025 AAGCCCCAGTGGCAGGACTCCAGACTCACTCTCCAGTGGGAGACAGAGATGTCCAGG 1084
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1023 aagccccagtgccagagactccagactcctccatctcagatgagagcagagatgtccag 1082
 OY 1085 GTACAGATGCAAGTNTTGTGAGGAATTGATGAGTCAAGATGATGAGTGTGAGAGCTGGG 1144
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1083 gtaacgatatcaagctctgtatgaggaactgtgacatgagatcagatgagatgagctgg 1142
 OY 1145 CTTNGCAGAGGAGTCTGGGAGACAGGAAGCAGATTTTCTGTGATT 1188
 ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1143 ctgagcagggagctctgggacagaagcagatcttctctgatt 1186

Search completed: September 22, 2002, 16:14:40
 Job time: 6835 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 14:49:28 : Search time 322.77 Seconds
(without alignments)
11809.333 Million cell updates/sec

Title: US-09-898-216-2

Perfect score: 1188

Sequence: 1 GGCTTCGGAGACNACGCT.....GGAAGCAGATTTCGTGATT 1188

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2138461 seqs, 1604250230 residues

Total number of hits satisfying chosen parameters: 4276922

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PCF_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq2:*
8: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1183	99.6	1188	5	US-09-898-216-2
2	1166	98.1	1322	6	US-10-219-793-106
3	1156.4	97.3	1910	6	US-10-198-846-12975
4	1145.8	96.4	1842	5	US-09-785-276A-24973
5	1144.4	96.3	1337	6	US-10-219-793-45
6	1130.4	95.2	1244	5	US-09-629-469A-11009
7	1104.8	93.0	9098	7	US-10-119-428-37
8	1086.4	91.4	1234	5	US-09-705-256A-3468
9	1057.4	89.0	1309	5	US-09-919-002-11863
10	815.6	68.7	1150	6	US-10-198-846-13409
11	618.6	52.1	622	1	PCF-US01-43704-1161
12	546.4	46.0	550	6	US-10-214-403-2995
13	538	45.3	563	1	PCF-US01-43704-58
14	516.6	43.5	691	7	US-10-097-105-1428
15	516.4	43.5	567	5	US-09-629-469A-1950
16	502	42.3	504	1	PCF-US01-43704-841
17	454	38.2	454	7	US-10-146-502-193
18	446	37.5	448	5	US-09-920-300A-1075
19	446	37.5	439	7	US-10-099-926-1075
20	423.4	35.6	439	1	PCF-US02-25766-1798
21	411.4	34.6	448	5	US-09-918-995-5335
22	394.4	33.2	475	5	US-09-918-995-5335
23	368.2	31.0	740	7	US-10-097-105-596
24	357	30.1	357	6	US-10-214-403-2873
25	352.8	29.7	362	5	US-09-721-544-8507

c	26	299.4	25.2	301	1	PCF-US02-12378-26	Sequence 26, Appl
c	27	299.4	25.2	301	7	US-10-124-805-26	Sequence 26, Appl
c	28	294	24.7	902	6	US-10-198-846-6734	Sequence 6734, Ap
c	29	292	24.6	300	5	US-09-539-331D-9558	Sequence 9558, Ap
c	30	289.4	24.4	521	1	PCF-US02-25766-8693	Sequence 8693, Ap
c	31	287	24.2	300	5	US-09-539-331D-21791	Sequence 21791, A
c	32	282.8	23.8	334	5	US-09-539-331D-28244	Sequence 28244, A
c	33	282.2	23.8	430	5	US-09-785-276A-37723	Sequence 37723, A
c	34	279.2	23.5	1407	5	US-09-886-497-17868	Sequence 17868, A
c	35	262	22.1	265	5	US-09-539-331D-14079	Sequence 14079, A
c	36	255	21.5	255	5	US-09-539-331D-4825	Sequence 4825, Ap
c	37	254.4	21.4	282	5	US-09-540-210B-26126	Sequence 26126, Ap
c	38	243.2	20.5	263	5	US-09-539-806B-2445	Sequence 2445, Ap
c	39	238	20.0	253	5	US-09-539-331D-688	Sequence 688, App
c	40	227.8	19.2	242	5	US-09-539-331D-10091	Sequence 30091, A
c	41	221	18.6	221	5	US-09-539-331D-15714	Sequence 15714, A
c	42	220	18.5	1123	6	US-10-219-999-8142	Sequence 8142, Ap
c	43	219	18.4	390	5	US-09-785-276A-7770	Sequence 7770, Ap
c	44	216.8	18.2	235	5	US-09-539-331D-29396	Sequence 29396, A
c	45	211	17.8	211	5	US-09-540-210B-18465	Sequence 18465, A

ALIGNMENTS

RESULT 1
US-09-898-216-2
Sequence 2, Application US/09898216
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
Goli, Surya K.
TITLE OF INVENTION: NOVEL HUMAN MEMBRANE PROTEIN
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Inocyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/898, 216
FILING DATE: 02-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/781,562
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0181 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1188 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE: Consensus
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-898-216-2

Query Match 99.68; Score 1183; DB 5; Length 1188;
Best Local Similarity 100.0%; Pred. No. 1,9e-289;
Matches 1188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTCTGGAGACNMCCGCTCCGCTGCTGTTGTTCCGAGAGTCCGCTGGCGGCTG 60
DB 1 GGCCTCTGGAGACNMCCGCTCCGCTGCTGTTGTTCCGAGAGTCCGCTGGCGGCTG 60
QY 61 GAAATGCTGGCG 120
DB 61 GAAATGCTGGCG 120
QY 121 GGCCTCTGGCG 180
DB 121 GGCCTCTGGCG 180
QY 181 TTCGTGCGCGACAGAGAGCGCTGGGTGGTGGAGCGCAATGGCGGATTCACCGGATCTG 240
DB 181 TTCGTGCGCGACAGAGAGCGCTGGGTGGTGGAGCGCAATGGCGGATTCACCGGATCTG 240
QY 241 GAGCCTGTTTGAACATCCTCATCCCTGTGTAGAACCGGATCCGATATGTGACAGATCTC 300
DB 241 GAGCCTGTTTGAACATCCTCATCCCTGTGTAGAACCGGATCCGATATGTGACAGATCTC 300
QY 301 AAGGAAATTTGATCATACAGTGGCTGAGCAGTGGCTGTGACTCTGCAATATTAATCTTG 360
DB 301 AAGGAAATTTGATCATACAGTGGCTGAGCAGTGGCTGTGACTCTGCAATATTAATCTTG 360
QY 361 CAAATGATGAGTGGCTTAACTGGCGCATGAGACCCCTTAAAGCAAGCAAGTACGGTGTG 420
DB 361 CAAATGATGAGTGGCTTAACTGGCGCATGAGACCCCTTAAAGCAAGCAAGTACGGTGTG 420
QY 421 GAGGACCCCTGAGTATGCCCTCACCCAGCTAGCTCAAAACCAATGAGATCAGAGCTCGGC 480
DB 421 GAGGACCCCTGAGTATGCCCTCACCCAGCTAGCTCAAAACCAATGAGATCAGAGCTCGGC 480
QY 481 AAACCTCTCTNTGGACAAATCTTCCGGGAACGGGAGTCCCTCAATGTCAGATTTGGGAT 540
DB 481 AAACCTCTCTNTGGACAAATCTTCCGGGAACGGGAGTCCCTCAATGTCAGATTTGGGAT 540
QY 541 GCCATCAACCAAGTCTGCTACTGCTGGGGTATCCGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 541 GCCATCAACCAAGTCTGCTACTGCTGGGGTATCCGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 601 ATCATGTGCGCAACCCCGGTGAAGTCTATGAGATGAGAGTGGAGGCGAGCGCGG 660
DB 601 ATCATGTGCGCAACCCCGGTGAAGTCTATGAGATGAGAGTGGAGGCGAGCGCGG 660
QY 661 AAACGGCGCCACAGTTCTAGAGTCTGAGGGGACCCGAGATGCCCATCAATGTGGCAGAA 720
DB 661 AAACGGCGCCACAGTTCTAGAGTCTGAGGGGACCCGAGATGCCCATCAATGTGGCAGAA 720
QY 721 GGGAGAAACAGAGCCAGATCTGCGCTCCGAAGCAGAAAGGCTAAAGCTAAGCTATTGCA 780
DB 721 GGGAGAAACAGAGCCAGATCTGCGCTCCGAAGCAGAAAGGCTAAGCTAAGCTATTGCA 780
QY 781 GCAGAGAGAGAGGCGCAGTCTGCGGAAAGGCCAAGGCTAAAGCTAAGCTATTGCA 840
DB 781 GCAGAGAGAGAGGCGCAGTCTGCGGAAAGGCCAAGGCTAAGCTAAGCTATTGCA 840
QY 841 ATCTGCTGCTGAGCTCTGACACAACTAATGAGATGAGAGCTTCACTGAGTGGCGC 900
DB 841 ATCTGCTGCTGAGCTCTGACACAACTAATGAGATGAGAGCTTCACTGAGTGGCGC 900
QY 901 GAGCAGTATGTAGGCGCTTCTCAAACTGGCAAGAGATCCAAACCTATCTCTAGTGGCC 960
DB 901 GAGCAGTATGTAGGCGCTTCTCAAACTGGCAAGAGATCCAAACCTATCTCTAGTGGCC 960
QY 961 TCCAAACCTGGCGATGTACACAGATGTGGCTCAGGCGCATGGGTGTATATGAGAGCCCTC 1020
DB 961 TCCAAACCTGGCGATGTACACAGATGTGGCTCAGGCGCATGGGTGTATATGAGAGCCCTC 1020
QY 1021 ACCAAAGCCCCAGTGGCAGAGGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGATGTC 1080

DB 1021 ACCAAAGCCCCAGTGGCAGAGGACTCCAGACTCACTCTCCAGTGGGAGCAGCAGATGTC 1080
QY 1081 CAGGCTACAGATGCAAGNTTGTATGAGAACTTGATCGAGTCAAGATGATGAGTGGAGC 1140
DB 1081 CAGGCTACAGATGCAAGNTTGTATGAGAACTTGATCGAGTCAAGATGATGAGTGGAGC 1140
QY 1141 TGGCTTGGCCAGGAGTCTGGGAGCAAGAAAGCAGATTTCTGATT 1188
DB 1141 TGGCTTGGCCAGGAGTCTGGGAGCAAGAAAGCAGATTTCTGATT 1188

RESULT 2
US-10-219-793-106
; Sequence 106, Application US/10219793
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; FILE REFERENCE: P2008P1C1
; CURRENT APPLICATION NUMBER: US/10/219,793
; PRIOR APPLICATION NUMBER: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/209,462
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: PCT/US98/12125
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/049,547
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,548
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,549
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,550
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,566
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,606
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,607
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,608
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,609
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,610
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,611
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/050,901
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/052,989
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058,665
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,668
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,669
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,750
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,971
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,972
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,975
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/060,834
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,841
; PRIOR FILING DATE: 1997-10-02

[illegible]

```

QY      1022 CCAAGGCCCACTGCGCAGGACTCCACGATCTACTCTCCAGTGGAGCGACAGATGTCC   1081
|         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |
Db      1022 ccaagcccccaagtgcacgaggactccagactcactctccagtgggcgccagagatgtccc   1081
QY      1082 AGGTCACAGATCAAGTTTGATGAGGAATTGATGCAGTCAAGATGAGTTAGTAGGAGCT   1141
|         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |
Db      1082 aggttcacgatgcaagttcttgaatgtaggaacttgatcgatgtaagaattgtagtgaact   1141
QY      1142 GGGCTTNGCCAGGAGCTCTGGGGACAAGAACCAGATTTCCTGATT   1188
|         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |
Db      1142 gggcttgccaggagctctggggacaagaaccagatttccctgaatt   1188

RESULT      4
US-09-785-276A-24973
; Sequence 24973, Application US/09785276A
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007B
; CURRENT APPLICATION NUMBER: US/09/785,276A
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24973
; LENGTH: 1842
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1841..1842
; OTHER INFORMATION: n = A,T,C or G
US-09-785-276A-24973

Query Match          96.4%; Score 1145.8; DB 5; Length 1842;
Best Local Similarity 99.2%; Pred. No. 5.4e-280;
Matches 1180; Conservative 0; Mismatches 7; Indels 3; Gaps 3

QY      1       1 GGGCTTGTGGAGGNACC-GCTCCGCTGTCGTTGTTGCCGAGAGTCCGTGCGCGGTG   59
|         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |
Db      479  gcctctgcgagcgacctcgcctcgtctcgttgttcgcgaagtgctgcgcgcgtg   538
QY      60       60 GGAAATGCTGCGCGCGCGCGCGCGGGGGGACACTGAGGCCCTTTTGGCTGAGAGGGGTCTCTAC   119
|         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |
Db      539  ggaaatgcttgcgcgcgcgcgcgcgcgcgcctccttcgcgcgaagggtcctaac   597
QY      120     120 TGACCTTGTGGCCCCGCGTCCG-CGCGGGCCGCTCTGGAATGCCCCGAACACCGTGGTAC   178
|         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |
Db      598  tggcttctgcgcgcgttcgcgcgcgcgcgcctccttcgcatggcccgaaaacacgtygtac   657
QY      179     179 TGTTCGTGCGCGACGAGAGGAGGCTGGGTGTGAGGCGAATGGCGCATTTCCACGGATTC   238
|         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |
Db      658  tgtctgtgccgacgagagagccttggtgtgtgagcgaatggcgaattccacccgatcc   717
QY      239     239 TGAAGCCTGGTTGAACATCTCATCCCTGTGTTAAGCCGGATCCGATATGTGACAGTCC   298

```


Dd 718 tggagccctgggttgaacatccatccctggtttagaccggatccgatatgycagaagtc 777
Qy 299 TCAGGAATTTGTCATCAACGTGCTGAGCAGTGGCTGTGACTCTCGCAATGTAATC 358
Dd 778 tcaaggaaattgtcatcaacagtcgctgagcagtcggtctgtgactcttcgacaatgttaacc 837
Qy 359 TCAGAAATGATGAGTGTCTTTTACCTCGCATCATGAGACCCCTTACAAGGCAAGCTACGGTG 418
Dd 838 tgcacaatcgatgagtccttaccctgcgcatcattgacaccttaacaggaagctacggtg 897
Qy 419 TGGAGGACCTGATGTCGCTGACCCAGCTGACCTCAACACACATGAGATGAGACTCG 478
Dd 898 tggaggaacccctgagtagtcgctgcaacccagctgagctcaaacacacatgagatcagaagctcg 957
Qy 479 GCAAACTCTCTTGTGCAAAAGTCTTCCGGGAACGGGAGTCCCTGAATGCAAGCATGTGCG 538
Dd 958 gcaaaactctctctgcaaaagctctccgggaacggagctccctgaatgcccagatctggtg 1017
Qy 539 ATGCCATCAACCAAGCTGCTGACTGCTGGGGTATCCGCTGCTTACGTTATGAGATCAAG 598
Dd 1018 atgcacataccaacagctgctgactgctggtgtatccgctgcctccgtttagatcaag 1077
Qy 599 ATATCCATGTCGACCCCGGGGGAAGAGTATGATGAGATGAGATGAGATGAGATGAG 658
Dd 1078 atatccatgctgcaaccccggtgaaagagctcatgagatgagatgagatgagatgagatgag 1137
Qy 659 GGAAGAGGCGCCACAGTCTGAGTGTGAGGGGACCCGAGAGTGGCGCATGATGTGCGAG 718
Dd 1138 ggaagagggccacagctctagagctgagggagagcccgagagctgagcacaatgctggtcg 1197
Qy 719 AAGGAGAAGAACAGCCAGATCTCTGCTCCGAGACAGAAAAAGCTGAACAGATTAATC 778
Dd 1198 aaggagaaagaaagagccagatccctgctcgaagcagaagaagctgagacagataatc 1257
Qy 779 AGGCACAGAGAGGCGCATGAGTCTGCGAAGGCCAAGGTTAAGTGAAGTGAAGTGAATC 838
Dd 1258 aggcagagaaagagagccagctgagcttctgagcgaagccaaagctaaagctaaagctaatc 1317
Qy 839 GAATCTGTGCTGAGCTGTGACACAAATATGAGATGAGATGAGATGAGATGAGATGAG 898
Dd 1318 gaatctgtgctgagctctgacacaaataatgagagatgagcagacttcaactgactggtg 1377
Qy 899 CCGAGCAGTATGTCAGCGGCTTCTCCAACTGCGCAAGAGTCTCAACACTATCTACTGC 958
Dd 1378 ccgagcagatgctcagcggtcttccaactgccaagagactccaacactatctactgc 1437
Qy 959 CCTCAACCTGCGCATGTCACACAGATGTCGTCAGGCCATGGTGTATATGAGAGCCC 1018
Dd 1438 cctcaaaccttgcagatgctaccacagcaaggctgagcagcatalggtgtatataggagccc 1497
Qy 1019 TCACCAAAAGCCAGTGCAGGAGCTCGAGACTCTCTCAGTGGGAGCAGCAGAGATG 1078
Dd 1498 tcaccaaaagcccaagctgccaagagactccaactctccagtgaggagcagcagagagtg 1557
Qy 1079 TCCAGGGTACAGATGCAAGTNTTGTATGAGAACTTGTATGAGTCAAGATGAGTTAGTGA 1138
Dd 1558 tccaggtacagatgcaagctctgtaggagaaacttgatcgagcagaagatgagtagtgga 1617
Qy 1139 GCTGGGCTTNGCCAGGAGTCTGGGGACAGACAGACAGATTTTCCGATT 1188
Dd 1618 gctgggcttngccagggagcttggggagcaagaagcagatttccctgatt 1667

RESULT 5
US-10-219-793-45

Sequence 45, Application US/10219793
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 86 Human Secreted Proteins
FILE REFERENCE: P2008P1C1
CURRENT APPLICATION NUMBER: US/10/219,793
CURRENT FILING DATE: 2002-08-16
PRIOR APPLICATION NUMBER: 09/209,462

PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: PCT/US98/12125
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/049,547
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,548
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,549
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,550
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,566
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,606
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,607
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,608
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,609
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,610
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,611
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/050,901
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/052,989
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/051,919
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/055,984
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/058,665
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,668
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,669
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,750
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,971
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,972
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,975
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/060,834
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,841
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,844
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,865
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/061,059
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/061,060
PRIOR FILING DATE: 1997-10-02
NUMBER OF SEQ ID NOS: 737
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 45
LENGTH: 1337
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (1335)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (1336)
OTHER INFORMATION: n equals a,t,g, or c

```
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1337)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-219-793-45
```

```
Query Match          96.3%: Score 1144.4; DB 6; Length 1337;
Best Local Similarity 99.4%: Pred. No. 1,1e-276;
Matches 1167; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

Oy 16 CCGCTCCGCTCTCTCTGTTGTTCCGAGATGCTGCGGGCTGGGAAATGCTGGCGGC 75
Db 28 ctgcctcgcctcctcgtctcgtccggaagtcgctgcgctggaatgctgcgcgc 87
Oy 76 GGGGGGGGGGGGACTGGGGGCCCCCTTTTGTGAGGGGCTCTACTAGGCTTCTGGCGCC 135
Db 88 ggggggc- ggggcaactggggcccttctgaggggtctctactctgcttcgcgcgc 146
Oy 136 TCCG- GCCGCGCTCTCTGATGATGGCCGAAACCGCTGACTTGTGTCGGCAGCA 194
Db 147 tccgcgccgcgcctcctcctcgtgattgcccggaaacacgctgctcgttcgcgcgca 206
Oy 195 GGAGCCCTGGGTGTTGAGACGAATGGCCGATTCACCCGATCTGAGACCTGGTTGAA 254
Db 207 ggaaggccctgggtgtgtgagcgaaatggccgattccacccggtatccctgagcttgaa 266
Oy 255 CATGCTCATCCCTGTGTTAGACCGGATCCGATATGTCGAGATGTCACGAATGTCAT 314
Db 267 catctccatccctgtgtatagaccggatccgatatgtgcagaatctcaaggaaatgtcat 326
Oy 315 CAACGTGCTGAGACAGTCGGCTGTGACTCTCGACAATGTAATGTAATGATGAGAGT 374
Db 327 caacgtgccttgagcagtcgctgcgtcctcgtacaaatgtaactctgcaatcgatgagct 386
Oy 375 CCTTACTGCGCCATCATGAGACCTTACAGCAAGCTACGGTGTGAGAGACCTGTAGTA 434
Db 387 ccttactcgcgcatacagacccttaccaggaagctacggtgtgaggaacccctgagta 446
Oy 435 TGCGCTCACCCAGTACGCTCAAAACAACATGATACAGAGTACGAGCAACTCTCTTTGA 494
Db 447 tgcgctcaaccgcagctcagctccaaacaacatgagatcagagctcgcgaaactctcttga 506
Oy 495 CAAAGTCTTCCGGGAACGGGAGTCCCTGATATGCCAGATTTGTGATGCCATCAACCA 554
Db 507 caaagtcttcccggaacgggagctccctgaaatgcagcatgtgagatgcatcaacaacagc 566
Oy 555 TGCTGACTGCTGGGGTATCCGCTGCTNGTATGATGATCAAGAGATATCCATGTGCCAC 614
Db 567 tgcgtacgtcgtgggtatctgcctgccttcaatgagatcaaggatatacctatgtgcacc 626
Oy 615 CCGGCTGAAGAAGTCTATGACAGATGACAGTGGAGGCGAGCGGGAACGGGCGCCAGT 674
Db 627 ccgggtgaaagagctctatgacagatgcaaggtgagagcagggcggaacgggcccagat 686
Oy 675 TCTAGAGTCTGAGGGGACCCGAGAGTCCGCATCAATGTGCGAGAAGGGAAGAAAGGC 734
Db 687 tctagagctctggggagaccggaagctcggtccatcaatgtgcagaagggaagaaacaagc 746
Oy 735 CCAGATCCGCGCTCCGAGAGCAAGAAAGGCTGACAGATTAATTCAGGCGAGGAGAGGC 794
Db 747 ccagatcccgctcccgagacagaaaggctgaaacagataaaatcaggcgaggaagagc 806
Oy 795 CAGTGCAGTTCTTGCGAGAGGCGCAAGGCTAAAGCTGAAGCTATTCGATCTGCTGCAC 854
Db 807 cagtgctagctcgcggaagggcaaggtcctaagctgaagctatctgaaatctcgtgcgacg 866
Oy 855 TCTGACACAACTAATGAGATGACAGACGCTTCACTGACTGTGCGCGAGCAGTATGTAC 914
Db 867 tctgacacaacaataatgagatgcagcagcttcaactgactgtgcgagcagatgctcag 926
Oy 915 GCGGTTTCCAAATGCGCCAAAGACTCCAAACTATCTACTGCGCTCCAAACCTGCGCA 974
Db 915 gggggtttccaaatgcgccaaagactccaaactatctactgcgctccaaacctgcgca 974
```

```
Db 927 cgcgtctcacaacatgcccagagatccacaactatccctactcgcctcccaacctgcga 986
Oy 975 TGTMCACAGCATGTGCGTCACAGCCATGGGTGATATGAGCGCTACCAAGCCCAT 1034
Db 987 tgcacacagatggtgcacagccatgggtgtatatagtagccctcaacaagccccagct 1046
Oy 1035 GCCAGGACTCCAGACTCACTCTCCAGTGGAGCAGACAGATGTTCCAGGATACAGATGC 1094
Db 1047 gccagagactccagactcactctcctcagtgaggcgacagatgctcaggtacagatgc 1106
Oy 1095 AAGTNTGATGAGGAACCTGATCAGTCAAGATGACTTGTGAGAGCTGGGCTTGGCCAG 1154
Db 1107 aagctcttgatgaggaactgactgacagatgaatgagctgagctgggcttgccag 1166
Oy 1155 GAGTCTGGGACAGAGACGATTTCTGTGATT 1188
Db 1167 gagctcgggacagaagagacagatcttcctgact 1200
```

RESULT 6

```
US-09-629-469A-11009
; Sequence 11009, Application US/09629469A
```

GENERAL INFORMATION:

```
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TENSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11009
; LENGTH: 1244
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35)..(1102)
```

```
US-09-629-469A-11009
```

```
Query Match          95.2%: Score 1130.4; DB 5; Length 1244;
```

```
Best Local Similarity 99.4%: Pred. No. 3.9e-276;
```

```
Matches 1153; Conservative 0; Mismatches 5; Indels 2; Gaps 2;
```

```
Oy 30 TCGTTGGTTCCGAGAGTCCCTGCGGCGTGGGAAATGCTGGCGCGCGCGGCGGCGCA 89
Db 1 tgcgtgttcgcgaggtcgtcgcggtggtggaatgctggcgcgcgcgcgcgcgcgcgcg 59
Oy 90 CTGGGCGCCCTTTTGTGAGAGGCTCTACTAGGCTTGTGCGCGCGCTCCG- GCCGCGCT 148
```

Db	60	ctggggcccttttgcttgaggggctctctactgctcttgcgcgcgcctccgcgcgcct	119
Oy	149	CTCTGCAATTGCCCCGGAACACCGTGTACTGTTCTGTCCGCGAGCAGAGGCTTGCGGTG	208
Db	120	ccctctgattgccccggaanaacccgtgtactcttgcgtccgcagcaggggctctgtgtg	179
Oy	209	TGGAGCGCAATGGGCGGATTTCACCGGATCCGGAGCTGGGTGTAACATCCATCCCTG	268
Db	180	tgagcgaaatgggcgcgtctccaccggaatccctggagcccttgatllgaacaccccaaccc	239
Oy	269	TGTTAGACCGGATCCGATATGTGCAGAGTCTCAAGCAAAATTGTCAATCAACGTGCTGAGC	328
Db	240	tgtttagaccggatccgatctatgtgcagagttccaaagaaatgtcatcaacgctctgcgc	299
Oy	329	AGTGGCGTGTGACTCTCGAATAATGTAACTCTGCAATTCATGAGTCTTTACTGTGCCA	388
Db	300	agtcggtgtgactctcgaacaatgtaactctgaaatcgatggaatctcttaccctgcga	359
Oy	389	TCATGAGACCCCTTACAAGCAAGCTACGGGTGTGGAGAACCCCTGAGTATGGCCGACACGAC	448
Db	360	tcatggaacctctaaagcgaagtaagtggtgagagacccttgagtatgctctcaaccgac	419
Oy	449	TAGCTCAAAACACCATGATGATGACAGCTCGGCAAACTCTCTNTGACAAAGTTCCGGG	508
Db	420	tagctcaaaacaaccatgatatgcagagctcgcgaacctctctctgtgacaagaagctctcgcg	479
Oy	509	AACGGGAGTCCCTCAATGCGCAGATGTGGATGCCATTCACCAACTGCTGACTGCTGGG	568
Db	480	aacgggagctccctgaatgacagcatctgtgatatgcatacaacaagctctgactgtcgtg	539
Oy	569	GTATCCCGTGCCTMCCTTATGAGATCAAGGATATCCATGTGCAACCCGGGTGAAAGAGT	628
Db	540	gtatccgctgcctcgttatgagatacaagatatcctatgtcaccccggtgtgaaagagt	599
Oy	629	CTATGCAGATGCAAGTGTGAGGCGAGCGCGGAACGGGCCACAGTTGTAGAGTGTGAGG	688
Db	600	ctatgcagatgcaagtggaagtcagagccgagcaacggtccaaagctctctajgtctcgtg	659
Oy	689	GGACCCGAGATCGGCCATCAATGTGGCGAAGAGGAAGAAAGGCCGAGATCCGCGCT	748
Db	660	ggaccgcgagagtcggccaatcaatgtgacgaaggaagaaacaagcccagatccctggcct	719
Oy	749	CCGAGACGAAAAAGCTGAACAGATAAATCGAGCAGCAGAGAGGCCAGTGCAGTTCTGG	808
Db	720	ccgaagcagaaagctggaacagataaatcaagcagcagagagagccagtgcaagttcgtg	779
Oy	809	CGAAGCGCAAGGCGTAAAGCTGAAGCTAATTCGATCTCTGGCTCCACTCTGCACACAACATA	868
Db	780	cgaaggtccaaagctcaaaagctgaaagctatctcgaaatccctgtgcagctctcgaaacaata	839
Oy	869	ATGAGATGCAAGCAAGCTTCACTACTACTGTGGCCGAGAGTATCTACAGCGGTCTTCCAAC	928
Db	840	atgagatgtgcagcagcttcaactgactgtgcgcgagagatgttcaagcggtcttccaac	899
Oy	929	TGGCCAGAGACTCCACACTATCTACTGCGCTCCAAACCTTGCGATGTACACAGCATGG	988
Db	900	tgggcaagagactccaacaactatctacttgcctcccaacccctgtgcgagtctaacagacgtg	959
Oy	989	TGGCTCAGGCGCATGGGCTGTATTATGGAGCCCTCACCAACCCCGAGTCCGAGGATCTTCAG	1048
Db	960	tggtcgaagccaatggtgtgtatatltgagggcccttcaccaagccccagctgcggyggaactcag	1019
Oy	1049	ACTCACTCTCCAGTGGGAGCAGCAGAGATGTCCAGGGTACAGATGCAAGTNTGTAGAG	1108
Db	1020	actcaactctccagctggagagcagcagagatgttccaggttacaagatgcaagctctgtatggg	1079
Oy	1109	AACCTTGACGATCAAGATAGTATTAGTGAGCGTGGGCTTNGCCAGAGAGTCTGGGGACAA	1168
Db	1080	aacttgatcgaggtccaagatgagatltagtgagctgcggcttgcgccaagagagctgtcgggacaa	1139
Oy	1169	GGAAGCAGATTTTCTGATT 1188	

```

DB      1140  ggaagcagatttcctgatt 1159

RESULT  9
US-10-119-428-37
: Sequence 37, Application US/10119428
: GENERAL INFORMATION:
: APPLICANT: Tang, Y. Tom
: APPLICANT: Liu, Chenghua
: APPLICANT: Asundi, Vinod
: APPLICANT: Xu, Chongjun
: APPLICANT: Wehrman, Tom
: APPLICANT: Ren, Feiyao
: APPLICANT: Ma, Yundong
: APPLICANT: Zhou, Ping
: APPLICANT: Zhao, Qing A.
: APPLICANT: Yang, Yonghong
: APPLICANT: Drmanac, Radjic T.
: TITLE OF INVENTION: Novel Nucleic Acids and
: TITLE OF INVENTION: Polypeptides
: FILE REFERENCE: 789CIP2
: CURRENT APPLICATION NUMBER: US/10/119,428
: CURRENT FILING DATE: 2002-04-09
: PRIOR APPLICATION NUMBER: 09/596,193
: PRIOR FILING DATE: 2000-06-17
: PRIOR APPLICATION NUMBER: 09/574,454
: PRIOR FILING DATE: 2000-05-19
: PRIOR APPLICATION NUMBER: 09/519,705
: PRIOR FILING DATE: 2000-03-07
: NUMBER OF SEQ ID NOS: 55
: SOFTWARE: PC_FL_genes Version 1.0
: SEQ ID NO 37
: TYPE: DNA
: LENGTH: 9098
: ORGANISM: Homo sapiens
: NAME/KEY: CDS
: LOCATION: (75)..(9098)
US-10-119-428-37

```

Query Match	93.0%;	Score.1104.8;	DB 7;	Length 9098;
Best Local Similarity	99.3%;	Pred. No. 1.7e-269;		
Matches 1128;	Conservative	0;	Mismatches 6;	Indels 2;
			Gaps	2

[illegible]

QY	420	GGAGGACCCCTGAGTATGGCCGTACCCAGCTAGCTAAACCAACATGAGATCAGAGCTGG	479
Db	431	ggaggagccctgagtagtcgcgttcaaccagctagcttaaaacaacatgagatcgaagctcgg	490
QY	480	CAAACTCTCTNTGCAAAAGTTCTCCGGGAACGGGAGTCCCGAATGCGACAGATTGTGGA	539
Db	491	caaaactctctcgtgcgaagaagctctccggtgaagggagctccctgaaatgcagcatgtgga	550
QY	540	TGCCATTCAACCAAGCTGCTGACTGCTGGGGTATCCGCTGCTTCGTTATGAGATCAAGGA	599
Db	551	tgcctatcaacaagatcgtcgtactgctcgtggtatccgcgtccgcttctatgatalcaagga	610
QY	600	TATCATGTGCGACCCCGGGTGAAGAGCTATGCGATGCGAGTGCAGGTGGAGCGACAGCGCG	659
Db	611	tatcatagtctgcaccccggtgcgaagaagctctatgaagatgcaggtgcgaagcgagcgcg	670
QY	660	GAACGGGCGCACAGTTCTAGAGTCTGAGGGGACCGCAGAGTGGCCATCAATTGTGGCAGA	719
Db	671	gaaaaggcgccacagctctcagagctcgtgagggagcccgagatcgcgcataatgtgcaga	730
QY	720	AGGGAGAAACAGGCGCGATCTCTGGCTCCGACACAAAAAGCTGAAACGATAAATCA	779
Db	731	agggaagaacaagcgcccgagatccctgcgtcccgaaacagaagaagctctgacagataaacatca	790
QY	780	GGCAGCAGAGAGGCGCAATGCGATTCTGTGCGAAGGCCAAGGCTAAAGCTGAAGCTATTTCG	839
Db	791	ggcagcagagagagcgagctgcagctctcgtcgtgagggccaaggtctaaagctgaagctatccg	850
QY	840	AATCCTGGCTCGAGCTGTGACACACATTAATGGAATATCAGACGCTTCACAGACTGTGGC	899
Db	851	aatcctgctcgtcgaagctctgcacacaacaataagagaatgcagaagctctcaatcgaactggtgc	910
QY	900	CGAGCAGTATGTACAGCGGCTTCTCCAAACTGGGCCAAGGACTTCAACACTATTCCTACTGCC	959
Db	911	cgaagcagatgtcagcgcggtctctccaactcgcgaagatctcaacactatctctactgcgc	970
QY	960	CTCCAAACCTCGGATGTACACAGCATGATGGTCTCAGGCCATGGGTGTATATGAGGCCCT	1019
Db	971	ctccaaacctcgtcgaatctgcacacagcaatggtcgtcgaagccatgaggtatataatgagccct	1030
QY	1020	CACCAAAAGCCCGACAGTGCAGGACACTGCAGACTCATCTTCCAGTGGGACAGCAGAGAGT	1079
Db	1031	caccaaaagcccgatgcgcagagatctcagactcactctccagttggtgcagcagagatggt	1090
QY	1080	CCAGGAGTACAGATGCAAGTNTTGATGAGGAACCTTATGATGAGTCAAGATGACTTAAT	1135
Db	1091	ccaggatcacagatgcaagctctgataaggaactctatcatgagatcaagaatgacttggct	1146

```

RESULT      8
US-09-705-256A-3468
; Sequence 3468, Application US/09705256A
; GENERAL INFORMATION:
; APPLICANT: Ma, Xiao-Dun; Dotson, Stanton B.; Monsanto Company
; TITLE OF INVENTION: Tumor Associated Molecules (TAMs); Targets for diagnosis, treatment
; TITLE OF INVENTION: and prevention of cancer
; FILE REFERENCE: 3214
; CURRENT APPLICATION NUMBER: US/09/705,256A
; CURRENT FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: US 60/164,285
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 8259
; SEQ ID NO 3468
; LENGTH: 1234
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-705-256A-3468

Query Match      91.4%; Score 1086.4; DB 5; Length 1234;
Best Local Similarity 99.2%; Pred. No. 5.4e-265;
Matches 1131; Conservative 0; Mismatches 5; Indels 4; Gaps 4;

```

QY	52	CGGCGGTGGGAAATGCTGGCGCGCGCGCGGGGCGACTGGAGCCCTTTTGTGTAGGGG	111
Db	1	cggcggctgggaaatgctggcgcgcgcgcgcg-99agagcattgggcccctttgtcgtagagg	59
QY	112	CTCTCTACTGGCCTTCGGCGCGGCTCGC--CGCGGCGCTCTCTGGATTGGCCCGAAAC	170
Db	60	ctcttactggcttcctggcgcgctcgcgcgcgcctcctcttgatctggcccgaataac	119
QY	171	CGTGGTACTGTTTCGTGGCGGAGAGAGGCGCT--GGGTGGTGGAGGAATGGGCCATTTCC	229
Db	120	cgltgtaactgtctgtcggcgagcagagagagcgctgggggtggtagagcgaatggcgatcc	179
QY	230	ACCGGATCTTGAGACCTGGTTTGAACATCTCAATCCCTGTGTAAACCGGATCCGATATG	289
Db	180	accggatctcggagccctggctgtgaactctcactccctgtgttagaccggatccgatatg	239
QY	290	TGCGAGATCTCAAGGAATTTGTATCATACGTGCTTACGATGTCGGCTGTGACTCTCGACA	349
Db	240	tgcgagatctcaaggaaattgtcaataaagcttgcgtgaagagctgggctgtgactctcgaca	299
QY	350	ATGTAACTCTCAATATGATGAGATGCTTTTACCTGGGATCATATGAGACCTTTCAGAGCA	409
Db	300	atgtaaactctgcaaatctgatgagagctcttactctgtgactcaatgaaacctcaaggcaa	359
QY	410	GCTACGGTGTGAGAGACCTTGAGTATGTCGTCACCCAGCTAGCTCAAACACATGAGAT	469
Db	360	gctaagggtgtgagagaccctggagatgctgcgtcaaccagcttagctcaaaacaacatgagat	419
QY	470	CAGAGCTGGGCAACCTCTGTTNTGGACAAATCTTTCGGGAGACGGGAGTCCCGAATGGCA	529
Db	420	cagagctgggcaaacctctctctgtgacaagaatctccgggaagcggagatccctgaaatgca	479
QY	530	GCATTGTGGATGGCCATCAACCAAGCTGTGACTGTCGTGGGGGTATCCGCTGCTTCATTATG	589
Db	480	gcattgtggatggccatcaaccaaagctgtgactgtcgtgggggtatccgctgcttcattatg	539
QY	590	AGATCAAGGATATTCATGTGTCACCCCGGTTAAGAGTCTTATGCAAGATGACAGTGGAGG	649
Db	540	agatcaaggatataccaatgtgtccaccgccgggtgaaagatctatgcagatgtcaggttagag	599
QY	650	CAGACCGCGGAAACGGGCGCACAGTTCTTAAGTCTGA--GGGACCGCGAGAGTCCGCCATC	708
Db	600	cagaccgcggaagaaagggccacagcttctaaagctcgaaggggaaacggagagtcggccatc	659
QY	709	AATGTGGCAGAGGAGGAAGAAACAGGCCCAATCTTGCCCTCCGAAGACAGAAAGCTCTAA	768
Db	660	aatgtggcagagagggaagaaacaggccccaatcttccgctccgaagacagaaagagctgaa	719
QY	769	CAGATTAATCAGSCAGCAGAGAGAGCCAGTGTGAGTTCTTGCGAAGGCCAAGGCTTAAAGCT	828
Db	720	cagataaatcagcagcagagagagagccagtgagctcttggcgaaggccaaagctaaagct	779
QY	829	GAAGCTAATTCGAATCTCGGCTGCGAGCTCTGTACACACATTAATGAGATGTGCACACTTCA	888
Db	780	gaagctatctgaatctccggctgagcctgtgacacaacataatgtgagatgtcagcagactca	839
QY	889	CTGACTGTGGCGGAGCAGTATGTACAGCGGCTTCTCCAAACTGGCCAAAGACATCCAAACT	948
Db	840	ctgactgtggcgcgagcagtatgtcagggcgcttctccaacatgtggccaagagactccaacct	899
QY	949	ATTCCTACTGCCCCCTCAACCCCTGGCGATGTACACAGCATGGTGGCTCAGGCCATGGGTGTA	1008
Db	900	atctcctactgcccctcaaaccccttggcgatgtcaacagatatgtgtctcaaggcaatgggtg	959
QY	1009	TATGGAGCCCTTCACCAAGGCCCAAGTGTCCAGGGACTTCACAGACTCTCCAGTGGAGGC	1068
Db	960	tatggagcccttcaccaaagcccaagctgtccagggatctccagactcatcttccagttggagc	1019
QY	1069	AGCAGAGATGTCCAGGGGTACACAGATGCAAGTNTTGATGAGAGAACTTGATGCAATCAAGTG	1128
Db	1020	agcagagatgtccaggggtacagatgtcaagatcttgtatgtgaagaaactgtatcgaatcaagc	1079
QY	1129	AGTTAGTGGACTGGGCTTNGCCAGGGAGTCTGGGACAAAGAACAGATTTTCTGTATT	1188

```
|||||
Db 1080 agtagtgagactgagctgagcgaagagctgagagacagaacagatttctgatt 1139

RESULT 9
US-09-919-002-11863/c
: Sequence 11863, Application US/09919002
: GENERAL INFORMATION:
: APPLICANT: Leshkowitz, Dena
: APPLICANT: Liu, Jin
: TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
: FILE REFERENCE: 20411-752CON1
: CURRENT APPLICATION NUMBER: US/09/919, 002
: CURRENT FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: APPLICATION NUMBER: 09/359, 922
: PRIOR FILING DATE: FILING DATE: 1999-07-22
: PRIOR APPLICATION NUMBER: APPLICATION NUMBER: US 09/034, 341
: PRIOR FILING DATE: FILING DATE: 1998-02-13
: NUMBER OF SEQ ID NOS: 13203
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 11863
: LENGTH: 1309
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-919-002-11863

Query Match 89.0%; Score 1057.8; DB 5; Length 1309;
Best Local Similarity 99.2%; Pred. No. 9.4e-258;
Matches 1092; Conservative 0; Mismatches 6; Indels 3; Gaps 3;

QY 91 TGGGGCCCTTTTGGCTGAGGGGCTCTACTGCGCTTGCGCCGCGCTCGG-GCGGGCGCTC 149
Db 1306 TGGGGCCCTTTTGGCTGAGGGGCTCTACTGCGCTTGCGCCGCGCTCGGCGCTC 1247

QY 150 CTCTGATTTGCCCGAACCCTGCTGCTGCTGCTGCGCGAGAGAGCCCT-GGATGG 208
Db 1246 CTCTGATTTGCCCGAACCCTGCTGCTGCTGCTGCGCGAGAGAGCCCTGGGGTGG 1187

QY 209 TGGAGCGCAATGGCGCGATTTCCACCGGATCTCTGGAGCTTGTTGAACATCTCATCCCTG 268
Db 1186 TGGAGCGCAATGGCGCGATTTCCACCGGATCTCTGGAGCTTGTTGAACATCTCATCCCTG 1127

QY 269 TGTAGACCGGATCCGATATGTGACAGATCTCAAGAAATTTGTCATCAACGCTCGAGAG 328
Db 1126 TGTAGACCGGATCCGATATGTGACAGATCTCAAGAAATTTGTCATCAACGCTCGAGAG 1067

QY 329 AGTCGGCTGTGACTCTGCACATGTACTGTGCAAAATGATGGAGTCTTACTGCGCA 388
Db 1066 AGTCGGCTGTGACTCTGCACATGTACTGTGCAAAATGATGGAGTCTTACTGCGCA 1007

QY 389 TCATGAGCCCTTACAAAGCAAGCTAGCGTGTGGAGAGCCCTGAGTATGCGCTACCCAGC 448
Db 1006 TCATGAGCCCTTACAAAGCAAGCTAGCGTGTGGAGAGCCCTGAGTATGCGCTACCCAGC 947

QY 449 TAGCTCAAAACACCATGATGATCAGAGTCTGGCAAACTCTCTTTGACAAAGTCTTCCGG 508
Db 946 TAGCTCAAAACACCATGATGATCAGAGTCTGGCAAACTCTCTTTGACAAAGTCTTCCGG 887

QY 509 AACGGAGATCCCTGAATCCAGCATTTGTGATGCCATCAACCAACCTCTGATCTCTGGG 568
Db 886 AACGGAGATCCCTGAATCCAGCATTTGTGATGCCATCAACCAACCTCTGATCTCTGGG 827

QY 569 GATCCGCTGCTGCTGATGATGATCAAGATATCCATGTC-CCACCCCG35TGAAGAAG 627
Db 826 GATCCGCTGCTGCTGATGATGATCAAGATATCCATGTC-CCACCCCG35TGAAGAAG 767

QY 628 TCTATGACAGATGACAGTGTGAGAGCAGAGCGGGAAGCGGCACAGTTCTTAGAGTCTGAG 687
Db 766 TCTATGACAGATGACAGTGTGAGAGCAGAGCGGGAAGCGGCACAGTTCTTAGAGTCTGAG 707

QY 688 GGGACCCGAGAGTGGCCATCATGTGCGACAGAAAGGAAACAGGCCCAAGATCTTGCGC 747
```

```
|||||
Db 706 GGGACCCGAGATTCGCCATCATATGTGGCAGAAAGGAAAGAACAGGCCACAGATTCCTGGCC 647

QY 748 TCCGAAGCAGAAAAGCGTGAACAGATTAATTCAGGCAGCAGAGAGGCCAGTCACTTCTG 807
Db 646 TCCGAAGCAGAAAAGCGTGAACAGATTAATTCAGGCAGCAGAGAGGCCAGTCACTTCTG 587

QY 808 GCGAAGGCCAAGCTTAAGCTTAAGCTTAATTCATCTGCTGCGAGCTCTTGACACAACT 867
Db 586 GCGAAGGCCAAGCTTAAGCTTAAGCTTAATTCATCTGCTGCGAGCTCTTGACACAACT 527

QY 868 AATGAGATGACAGACGCTTCACTGAGTGTGGCCGAGCAGATATGTGAGCGCGTTCTCCAA 927
Db 526 AATGAGATGACAGACGCTTCACTGAGTGTGGCCGAGCAGATATGTGAGCGCGTTCTCCAA 467

QY 928 CTGGCCCAAGACTCCACACATCTCTACTGCGCTCCACACCTGGCGATGTCAACGACATG 987
Db 466 CTGGCCCAAGACTCCACACATCTCTACTGCGCTCCACACCTGGCGATGTCAACGACATG 407

QY 988 CTGGCTCAGAGCCATGGGTATATGAGAGCCCTCACCAGAACCCCAAGTGCAGGAGCTCCA 1047
Db 406 CTGGCTCAGAGCCATGGGTATATGAGAGCCCTCACCAGAACCCCAAGTGCAGGAGCTCCA 347

QY 1048 GACTCACTCTCCAGTGGGAGCAGACAGATGTCCAGGGTACAGATGCAATTTGATGAG 1107
Db 346 GACTCACTCTCCAGTGGGAGCAGACAGATGTCCAGGGTACAGATGCAATTTGATGAG 287

QY 1108 GAACCTTGATCGATCAAGATGCTTAGTGTGAGCTGGCGCTTCCACAGGAGTCTGGGACA 1167
Db 286 GAACCTTGATCGATCAAGATGCTTAGTGTGAGCTGGCGCTTCCACAGGAGTCTGGGACA 227

QY 1168 AGAAGCAGATTTCTGATT 1188
Db 226 AGAAGCAGATTTCTGATT 206
```

```
RESULT 10
US-10-198-846-13409
: Sequence 13409, Application US/10198846
: GENERAL INFORMATION:
: APPLICANT: Lillie, James
: APPLICANT: Xu, Yongyao
: APPLICANT: Wang, Youzhen
: APPLICANT: Steinhmann, Kathleen
: TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
: TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
: FILE REFERENCE: MRI-049
: CURRENT APPLICATION NUMBER: US/10/198, 846
: CURRENT FILING DATE: 2002-07-18
: PRIOR APPLICATION NUMBER: 60/306, 220
: PRIOR FILING DATE: 2001-07-18
: NUMBER OF SEQ ID NOS: 14084
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 13409
: LENGTH: 1150
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc-feature
: LOCATION: 1149, 1150
: OTHER INFORMATION: n = A,T,C or G
US-10-198-846-13409
```

```
Query Match 68.7%; Score 815.6; DB 6; Length 1150;
Best Local Similarity 97.9%; Pred. No. 1.8e-196;
Matches 856; Conservative 0; Mismatches 13; Indels 5; Gaps 3;

QY 318 CGTGCTAGCAGCTGGCGTGTGACTCTGACAAATGTAATCTGCAAAATGATGAGTCTCT 377
Db 107 cgggctcagcagctgctgctgctcctcgacaaatgtaactctgcaaatcagctcctcct 166
```

```
QY 378 TTACCTGCGCATCATGACCCCTTACAAAGGCAAGCTACGGTGTGGAGACCCCTGAGTATGC 437
D 167 ttaccctgcacatacctgagcccttaacagcaagctacggtg tggagagcccttgatgc 226
QY 438 CGTACCACGACTACTCAACACCATGAGATCAGAGCTGGGCAACCTCTGTTGAGACAA 497
D 227 cgtaccacagctacatacacaacacatgagctcagagctcggcaaacctctctcggacaa 286
QY 498 AGTCTTCGGGAGGAGGAGTCCCTGAATGTCAGCATTTGTGATGTCATCAACCAAGCTGC 557
D 287 agcttcctgggaagaggggtctccctgaatcagcatctg tggatgcatcaacccaagctgc 346
QY 558 TGAATGCTGGGGTATCCGCTCTGCTGCTGTTATGATGATCAAGATATCCATGTCACCCGC 617
D 347 tgaatgctgggtatacctgcctctgctctgatacgaatcaagatacctcgtcgcaccccg 406
QY 618 GGTAAAGAGCTATGATGATGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 677
D 407 ggtaaagagctatacctgcagatgagagctgagagcagagcagagagagagagagagagctc 466
QY 678 AGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 737
D 467 agagctgagagagagagagagagagagagagagagagagagagagagagagagagagagagag 526
QY 738 GATCTGCGCTCCGACAGCAAGAAAGCTGAAACAGATTAATAGCAGCAGCAGCAGCAGCAG 797
D 527 gatctgctgctccgacagcaagaaagctgaaacagattaatagcagcagcagcagcagcag 586
QY 798 TGCAGTTCTGGCGAAGGCGTAAAGCTGAAAGCTATGGAATCCGTGGTGGAGGAGTCT 857
D 587 tgcagttctggcgaaggcgtaaaagctgaaagctatggaaatccgtggtagagctc 646
QY 858 GACACAACATTAATGAGATGAGAGCAGCTTCACTGACTGTGGCCGAGCAGAGTATGTCAGCGC 917
D 647 gacacaacataatgagatgagagcagcttcaactgactgtggccgagcagagatg tgcagcgc 706
QY 918 GTTCTCCAACTGGCCCAAGGACTCCACACTATCTTACTGCTCCCTCCAACTGGCGAGTGT 977
D 707 gttctccaaactggcccaaggactccacactatcttactgctccctccaaactggcgaagt 766
QY 978 CACCAAGCATGTGGCTCAGGCGATGGGTATATGAGACCCCTCAACCAAGGCCAGTGGCC 1037
D 767 caccagcatgtggctcagcgcatgggtatatgagacccctcaaccaaaggccagtgcc 826
QY 1038 AGGAGCTCCAGACTCAGCTCTTCAGTGGAGCAGCAGAGATGTCAGAGGTACAGATGCAAG 1097
D 827 aggagctccagactcagctcttcagtgaggagcagagatgtcca -ggtacagatgcaag 884
QY 1098 TTTTGATGAGAGAACTTGATCGATCGATGAGATGAGTTAGTGAGCTGGG--CTTGCCAGG 1154
D 885 ttttgatgagagaaacttgatcgatcgatgagatgagttagtgagctggagcttgagccag 944
QY 1155 GAGTCTGGGAGCAAGAAAGATTTTCTGATT 1188
D 945 gagtctgggagcaagaaagatlttctccgatt 978
```

```
RESULT 11
PCT-US01-43704-1161/C
: Sequence 1161, Application PC/JTUS0143704
: GENERAL INFORMATION:
: APPLICANT: Corixa Corporation
: APPLICANT: Stolx, John A.
: APPLICANT: Xu, Jiangchun
: APPLICANT: Chenault, Ruth A.
: APPLICANT: Meagher, Madelein Joy
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
: FILE REFERENCE: 210121.561PC
: CURRENT APPLICATION NUMBER: PCT/US01/43704
: NUMBER OF SEQ. ID NOS: 2606
: SOFTWARE: Corixa Invention Disclosure Database
```

```
: SEQ. ID NO 1161
: LENGTH: 622
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 615
: OTHER INFORMATION: n = A,T,C or G
PCT-US01-43704-1161
```

```
Query Match 52.1% Score 618.6; DB 1; Length 622;
Best Local Similarity 99.4% Pred. No. 1,1e-146;
Matches 618; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 310 GTCATCAACGTCGTCGACGACTGCGCTGTGACTCTCCGACATGTAATCAATGAT 369
D 622 GTCATCAACGTCGTCGACGACTGCGCTGTGACTCTCCGACATGTAATCAATGAT 563
QY 370 GGAGTCTCTTACCTGCGCATGATGACCCCTTACAAGGCAAGCTACGCTGTGAGAGCCCT 429
D 562 GGAGTCTCTTACCTGCGCATGATGACCCCTTACAAGGCAAGCTACGCTGTGAGAGCCCT 503
QY 430 GAGTATGCGCTCACCCAGCTAGCTCTCAACAAACATGAGATAGAGCTTGGCAACTCTCT 489
D 502 GAGTATGCGCTCACCCAGCTAGCTCTCAACAAACATGAGATAGAGCTTGGCAACTCTCT 443
QY 490 NTGCAAAAGCTCTCCGGGAGCGGAGTCCCTGAATGCGCAGCATTTGTGATGCCATCAAC 549
D 442 CTGACAAAGCTCTCCGGGAGCGGAGTCCCTGAATGCGCAGCATTTGTGATGCCATCAAC 383
QY 550 CAAGCTGCTGACTGCTGGGATATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 609
D 382 CAAGCTGCTGACTGCTGGGATATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 323
QY 610 CCACCCCGGGTGAAGAAGTCTATGACAGTGCAGTGCAGGCGCAGAGCGCGCAACCGGCC 669
D 322 CCACCCCGGGTGAAGAAGTCTATGACAGTGCAGTGCAGGCGCAGAGCGCGCAACCGGCC 263
QY 670 ACAGTTCTAGAGTCTGAGGAGACCGGAGTCCGACATGATGATGAGGAGGAGGAGGAGGAG 729
D 262 ACAGTTCTAGAGTCTGAGGAGACCGGAGTCCGACATGATGATGAGGAGGAGGAGGAGGAG 203
QY 730 CAGGCCAGATCTGCGCTCCGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 789
D 202 CAGGCCAGATCTGCGCTCCGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 143
QY 790 GAGGCCAATGAGTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 849
D 142 GAGGCCAATGAGTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 83
QY 850 GCAGCTCGACACAAACATTAATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 909
D 82 GCAGCTCGACACAAACATTAATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 23
QY 910 GTCAGCGCGTCTCCAAACTGG 931
D 22 GTCAGCGCGTCTCCAAACTGG 1
```

```
RESULT 12
US-10-214-403-2995/C
: Sequence 2995, Application US/10214403
: GENERAL INFORMATION:
: APPLICANT: Jiang, Yugu
: APPLICANT: Chenault, Ruth A.
: APPLICANT: Xu, Jiangchun
: APPLICANT: Indrias, Carol.Yoseph
: APPLICANT: Lodes, Michael J.
: APPLICANT: Secrist, Heather
: APPLICANT: Carter, Darrick
: APPLICANT: Fanger, Gary R.
: APPLICANT: Smith, Carole L.
```



```

: CURRENT FILING DATE: 2002-03-13
: NUMBER OF SEQ ID NOS: 1562
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1428
: LENGTH: 691
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: 439, 508, 523, 526, 539, 561, 575, 582, 584, 602, 614, 623,
: LOCATION: 626, 645, 648, 664
: OTHER INFORMATION: n = A,T,C or G
US-10-097-105-1428
```

```

Query Match          43.5%: Score 516.6; DB 7; Length 691;
Best Local Similarity 91.4%: Pred. No. 7,4e-121;
Matches 616; Conservative 0; Mismatches 46; Indels 12; Gaps 7;
```

```

QY 200 CCTGGTGTGGAGCGAATGGCCGATTCCACCGGATCCTGAGCCTGGTTTGAACATCC 259
Db 1 cctgggtgtggagcgaaatggccgattccacccggatcctggagcctgtgtgaacatcc 60
QY 260 TCATCCCTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAGAGAAATTGTTCATCAACG 319
Db 61 tcattccctgtgtagaccggatccgatatgtgcagagctccaagaaattgtcatcaacg 120
QY 320 TGCCGTGACAGTCGGCTGTGACTCTGACAAATGTACTCTGCAATCGATGGAGTCTTTT 379
Db 121 tgccgtgacagtcggctgtgactctgacaaatgtactctgcaaatcgatggagtccttt 180
QY 380 ACCTGCGCATATGAGACCCCTTACAAGCAGCTACGCTGTGGAGAGCCCTGAGTATGCCG 439
Db 181 acctgcgcatatgagaccccttacaagcagctacgctgtggagagccctgagatgccc 240
QY 440 TCACCCAGCTAGCTTAAACCAACATGAGATCAGACCTGGCAAACTCTCTNTGACAAAG 499
Db 241 tcacccagctagcttaaaccaacatgagatcagacctggcaaaactctctntgacaaag 300
QY 500 TCTTCCGCGAAGCGGAGTCCCTGAAATGCCAGCATTTGATGCCATCAACCAAGCTGCTG 559
Db 301 tcttccgsgaagcgagtccttgaaatgccagcatatgagcctcaacaaagctgctg 360
QY 560 ACTGCTGGGGTATCCGCTGCTTCTGATGAGATCA--GATATTCATGTGCCACCCGG 618
Db 361 actgctggggtatccgctgcttctgatgagatcaaggaatcattgtgccaccgg 420
QY 619 GTGAAG--AGTCTATGAGATGAGTGGAGGACAGC--GGCGGAAAGGGCCACAGTTC 676
Db 421 gtgaag--agtctatgagatgagtgagggagcagagcgaggaacggccacagttc 480
QY 677 TAGAGTCTGAGGAGCAGCCGAGAGTGGCCATCAATGTGCGCAGAAAG---GAAGAAACA 731
Db 481 tagagtcctgagggagccggaagtcgncatcattgtgsgaanaagggaagaacaang 540
QY 732 GGCCAGATCTCTGGGCT--CCGAAGCAGAAAAAGCTGAACAGA--TAATCAGAGCAGAG 788
Db 541 ggcacgaattctctgggcttccgaagaaagatgaacaaantaaatcaagccacgg 600
QY 789 AG-AGCCAGTACTTGTGCGGAGCCAGAGGCTAAAGCTAATGTAAGCTATTCATCTGG 847
Db 601 aagagccatctgctgttcttgcaanaagcgaaggtcaaaactgaanctntcgaaatcc 660
QY 848 CTCAGCTCTGACA 861
Db 661 gtgnaactttacca 674
```

```

RESULT 15
US-09-629-469A-1950
: Sequence 1950, Application US/09629469A
: GENERAL INFORMATION:
: APPLICANT: OTA, TOSHIO
```

```

: APPLICANT: ISOGAI, TAKAO
: APPLICANT: NISHIKAWA, TETSUO
: APPLICANT: HAYASHI, KOJI
: APPLICANT: SATTO, KAORU
: APPLICANT: YAMAMOTO, JUNICHI
: APPLICANT: ISHII, SHIZUKO
: APPLICANT: SUGIYAMA, TOMOYASU
: APPLICANT: WAKAMATSU, AI
: APPLICANT: NAGAI, KEIICHI
: APPLICANT: OTSUKI, TETSUJI
: TITLE OR INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
: FILE REFERENCE: 084335/0123
: CURRENT FILING DATE: US/09/629, 469A
: PRIOR FILING DATE: 2000-07-28
: PRIOR APPLICATION NUMBER: JP 1999-248036
: PRIOR FILING DATE: 1999-07-29
: PRIOR APPLICATION NUMBER: JP 1999-300253
: PRIOR FILING DATE: 1999-08-27
: PRIOR APPLICATION NUMBER: JP 2000-118776
: PRIOR FILING DATE: 2000-01-11
: PRIOR APPLICATION NUMBER: JP 2000-183767
: PRIOR FILING DATE: 2000-05-02
: PRIOR APPLICATION NUMBER: JP 2000-241899
: PRIOR FILING DATE: 2000-06-09
: PRIOR APPLICATION NUMBER: 60/159,590
: PRIOR FILING DATE: 1999-10-18
: PRIOR APPLICATION NUMBER: 60/183,322
: PRIOR FILING DATE: 2000-02-17
: NUMBER OF SEQ ID NOS: 19025
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1950
: LENGTH: 567
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (258)..(258)
: OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (333)..(333)
: OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (379)..(379)
: OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (426)..(426)
: OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (438)..(438)
: OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (441)..(441)
: OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (445)..(445)
: OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
: NAME/KEY: misc.feature
: LOCATION: (508)..(508)
: OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
: NAME/KEY: misc.feature
: LOCATION: (533)..(533)
: OTHER INFORMATION: "n" may be a, t, c, g, other or unknown
: NAME/KEY: misc.feature
: LOCATION: (561)..(561)
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 13:21:14 : Search time 1684.32 seconds
(without alignments)
9519.796 Million cell updates/sec

Title: US-09-898-216-2

Perfect score: 1188
Sequence: 1 GGCCTCTGCGAGCACCCT.....GGAGCAGATTTCCTGATT 1188

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27473414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_huv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1067.2	89.8	1218	11 AF161458	AF161458 Homo sapi
2	916.6	77.2	935	9 AL541792	AL541792 AL541792
3	899.4	75.7	1110	10 BM476304	BM476304 AGENCOURT
4	884.2	74.4	962	9 AL516503	AL516503 AL516503
5	871.8	73.4	994	10 BM468527	BM468527 AGENCOURT
6	860	72.4	998	10 BM468407	BM468407 AGENCOURT
7	842.2	70.9	1536	11 AK002793	AK002793 Mus muscu
8	835.4	70.3	1011	10 BM461614	BM461614 AGENCOURT
9	821.4	69.1	849	9 AL525168	AL525168 AGENCOURT
10	821	69.1	1061	10 BM453159	BM453159 AGENCOURT
11	802.6	67.6	899	9 AL525167	AL525167 AL525167
12	792.6	66.7	1095	10 BM423347	BM423347 AGENCOURT
13	791.8	66.6	918	10 BM459766	BM459766 AGENCOURT
14	791.4	66.6	943	10 BM479080	BM479080 602526159
15	791.2	66.6	920	10 B1862105	B1862105 60330995
16	791	66.6	925	10 BM043426	BM043426 603619796
17	790.8	66.6	820	10 BG685336	BG685336 602637769

C	18	770.4	64.8	787	9	AL569086
	19	766.2	64.5	796	10	B1092184
	20	764.4	64.3	1058	10	BM423453
	21	762.8	64.2	1053	10	BM456748
	22	757.4	63.8	864	10	BM765611
	23	757.4	63.8	887	10	BM045426
	24	755.6	63.6	923	10	B1870704
	25	752.4	63.3	853	10	BE314702
	26	750.4	63.2	988	10	BM048505
	27	749	63.0	830	10	BC575196
	28	743.4	62.6	836	10	BG823189
	29	740.8	62.4	924	10	B1767424
	30	738	62.1	1017	10	BG753777
	31	736.2	62.0	778	10	BG681190
	32	724.6	61.0	809	10	B1086063
C	33	716.8	60.3	770	10	BG828070
	34	715.4	60.2	903	10	BG532231
	35	711.6	59.9	813	10	BG768715
	36	709.4	59.7	856	10	BF965844
	37	706.2	59.4	916	10	B1260046
	38	703.8	59.2	752	10	B1552983
	39	702	59.1	728	10	B1522493
	40	702	59.1	739	10	BG574270
	41	701.2	59.0	865	10	BM051182
	42	700	58.9	986	10	B1694744
	43	698.4	58.8	788	10	BG767007
	44	698.2	58.8	785	10	BM009759
	45	697.4	58.7	1190	10	BF026777

ALIGNMENTS

RESULT	1
AF161458	AF161458
LOCUS	Homo sapiens HSPC108 mRNA, complete cds.
DEFINITION	AF161458
ACCESSION	AF161458.1
VERSION	GI:6841439
KEYWORDS	HTC.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Zhang, Q.H., Ye, M., Ren, S.X., Zhao, M., Zhao, C.J., Fu, G., Shen, Y., Fan, H.Y., Lu, G., Zhong, M., Xu, X.R., Han, Z.G., Zhang, J.W., Tao, J., Huang, Q.H., Zhou, J., Hu, G.X., Gu, J., Chen, S.J. and Chen, Z.
TITLE	Cloning and functional analysis of cDNAs with open reading frames for 300 previously undefined genes expressed in CD34+ hematopoietic stem/progenitor cells
JOURNAL	Genome Res. 10 (10), 1546-1560 (2000)
MEDLINE	20499367
PUBMED	11042152
REFERENCE	2 (bases 1 to 1218)
AUTHORS	Ye, M., Zhang, Q.H., Zhou, J., Shen, Y., Wu, X.Y., Guan, Z.O., Wang, L., Fan, H.Y., Mao, Y.F., Dai, M., Huang, Q.H., Chen, S.J. and Chen, Z.
TITLE	Human full length cDNA cloned from cd34+ stem cells
JOURNAL	Unpublished
AUTHORS	3 (bases 1 to 1218)
TITLE	Direct Submission
JOURNAL	Submitted (23-JUN-1999) Shanghai Institute of Hematology, Shanghai Second Medical University, Rui-Jin Hospital, 197 Rui-Jin Road II, Shanghai 200025, P. R. China.
FEATURES	location/Qualifiers
source	1..1218
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="CBCAPB12"
	/cell_type="cd34+ stem cells"
	/tissue_type="blood"

```

COS
18. 1046
/codon_start=1
/product="HSPc108"
/protein_id="AAF29073.1"
/db_xref="GI:6841440"
/translation="MRGSLASGRACRASSGLPENTVVLFPVPOEAMVVERMREHRI
LEPRLNLIPLVDRIQRTVRSLSKEVINVPDPSATLINDVLIYLRIMDPKAS
YVEDPEYATVQLQRTVRSLSKEVINVPDPSATLINDVLIYLRIMDPKAS
YEIDIVHPRVRKESKQMOYEAERRRKPTVLESGTRESAINVNEGKQADVLSAE
KAEQINOAGPASAIVLAKAKAKAEAIRILAAALOHNGDDAASLTVAEOVYASSLA
KDSNTILLPSNPGBVTSMVAVQAMGVYALTRKAPVGPIDPSLSSSSSDVOGTDSLDE
ELDRVKS"
BASE COUNT      313 a      310 c      337 g      258 t
ORIGIN
Query Match      89.8%; Score 1067.2; DB 11; Length 1218;
Best Local Similarity 99.1%; Pred. No. 1.2e-250;
Matches 1092; Conservative 0; Mismatches 7; Indels 3; Gaps 2;

QY 88 CACTGGGGGCCCCCTTTTGC-TGAGGGGCTCTTACTGGCTTTCGGCCGCGCCGCGC 146
Db 1 CACTGGGGGCCCCCTTTTGCATGAGGGGCTCTTACTGGCTTTCGGCCGCGCTT--GCCGCGC 58

QY 147 CTCCTCTGATTGCCCCCAACACCGTGGTACTGTCGCGCGCAGCAGAGGCGCGGT 206
Db 59 CTCCTCTGATTGCCCCCAACACCGTGGTACTGTCGCGCGCAGCAGAGGCGCGGT 118

QY 207 GGTGAGCGAATGGGCGCATTCACCGGATCCTGGAGCGCTGTTGACATCCTCATCC 266
Db 119 GGTGAGCGAATGGGCGCATTCACCGGATCCTGGAGCGCTGTTGACATCCTCATCC 178

QY 267 TGTGTAGACCGGATTCGATATGTGCAAGTCTCAAGAAATTGTCATCAAGTGCCTTA 326
Db 179 TGTGTAGACCGGATTCGATATGTGCAAGTCTCAAGAAATTGTCATCAAGTGCCTTA 238

QY 327 GGAGTCGGTGTGACTCTGCACATGTACTGCAAAATCGATGGATGCTTACTGCG 386
Db 239 GGAGTCGGTGTGACTCTGCACATGTACTGCAAAATCGATGGATGCTTACTGCG 298

QY 387 CATCATGACCCCTTACAGGACGATACGCTGTGAGAGCCCTGAGTATCCGTCACCA 446
Db 299 CATCATGACCCCTTACAGGACGATACGCTGTGAGAGCCCTGAGTATCCGTCACCA 358

QY 447 GCTAGCTCAACACCATGATGAGTCCGCAAACTCTCTNTGACAAAGTCTTCG 506
Db 359 GCTAGCTCAACACCATGATGAGTCCGCAAACTCTCTCTGACAAAGTCTTCG 418

QY 507 GGAACGGGATCCCTGAAATGCCAGATGTCATCAACCAAGCTGCTGACTGCTG 566
Db 419 GGAACGGGATCCCTGAAATGCCAGATGTCATCAACCAAGCTGCTGACTGCTG 478

QY 567 GGGTATCCCTGCTCCGCTTATGATCAAGATATCATGTCGCAACCCGCGGTGAAGA 626
Db 479 GGGTATCCCTGCTCCGCTTATGATCAAGATATCATGTCGCAACCCGCGGTGAAGA 538

QY 627 GTCTATGCAAGTCAAGTGGAGGCGGCGGAAACGGGCCACAGTTCTAGAGTCTGA 666
Db 539 GTCTATGCAAGTCAAGTGGAGGCGGCGGAAACGGGCCACAGTTCTAGAGTCTGA 598

QY 687 GGGGACCCCAAGATCCGCCATCAATGTGCGAGAAGGAAGAAACAGGCCAGATCCGCGC 746
Db 599 GGGGACCCCAAGATCCGCCATCAATGTGCGAGAAGGAAGAAACAGGCCAGATCCGCGC 658

QY 747 CTCGGAAGCAGAAAGGCTGACAGATTAATCAAGCAGCAGAGAGGCGCAGTCTT 806
Db 659 CTCGGAAGCAGAAAGGCTGACAGATTAATCAAGCAGCAGAGAGGCGCAGTCTT 718

QY 807 GGGGAAGGCGCAAGGCTAAAGCTATTCGAATCTGCGTGCAGCTTGACACACA 866
Db 719 GGGGAAGGCGCAAGGCTAAAGCTATTCGAATCTGCGTGCAGCTTGACACACA 778

QY 867 TAATGAGATGCAGCAGCTTCACTGACTGTGGCGAGCAATATGTCAAGCGCGTTCTCAA 926

```

```

Db 779 TAATGAGATGCAGCAGCTTCACTGACTGTGGCGAGAGATATGCAGCGGCTTCCAA 838
QY 927 ACTGGCCAAAGACATCCAAACATATCTTACTGCGCTCCAAACCTGGCGATGTCACCAGAT 986
Db 839 ACTGGCCAAAGACATCCAAACATATCTTACTGCGCTCCAAACCTGGCGATGTCACCAGAT 898
QY 987 GGTGGCTCAGGCGATGGGTGATATGAGAGCCCTTCACAAAGCCCGAGGCGGAGACTCC 1046
Db 899 GTGGCTCAGGCGATGGGTGATATGAGAGCCCTTCACAAAGCCCGAGGCGGAGACTCC 958
QY 1047 AGACTCACTCTCAGTGGAGCAGCAGAGATGTCAGAGGTACAGATGCAAGTNTTGATGA 1106
Db 959 AGACTCACTCTCAGTGGAGCAGCAGAGATGTCAGAGGTACAGATGCAAGTNTTGATGA 1018
QY 1107 GGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1166
Db 1019 GGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1078
QY 1167 AAGGAAGCAGATTTTCCTGATT 1188
Db 1079 AAGGAAGCAGATTTTCCTGATT 1100

RESULT 2
LOCUS AL541792 935 bp mRNA linear EST 16-FEB-2001
DEFINITION AL541792 LTI_FL002_PL1 Homo sapiens CDNA clone CS0DE007YE18 5 prime
, mRNA sequence.
ACCESSION AL541792
VERSION AL541792.1 GI:12873201
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 935)
AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..935
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DE007YE18"
/clone_1ib="LTI_FL002_PL1"
/lab_host="DH10B"
/note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand
cDNA was primed with a Notti-oligo(dT) primer. Five prime
end enriched, double-stranded cDNA was digested with Not I
and cloned into the Not I and Eco RV sites of the
pCMVSPORT 6 vector. Library was constructed by life
technologies. Contact : Feng Liang Life Technologies, a
division of invitrogen 9800 Medical Center Drive Rockville
, Maryland 20850, USA fax : (1) 301 610 8371 Email :
filiang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 239 a 246 c 265 g 181 t 4 others
ORIGIN
Query Match 77.2%; Score 916.6; DB 9; Length 935;
Best Local Similarity 99.0%; Pred. No. 7.9e-214;
Matches 927; Conservative 3; Mismatches 5; Indels 1; Gaps 1;
Db 189 GCAGCAGAGAGCGCTGGTGTGAGAGCAATGGGCCATTTCCACCGGATCTGAGGCTGG 248
Db 1 GCAGCAGAGAGCGCTGGTGTGAGAGCAATGGGCCGATTCACCGGATCTGAGGCTGG 60

```

```
QY 249 TTTGAACATCTCATCCCTGTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGAAAT 308
|||||
Db 61 TTGGAACATCTCATCCCTGTGTGTAGACCGGATCCGATATGTGCAGAGTCTCAAGAAAT 120
QY 309 TGTTCATCAACGTGCTGAGCAGTCCGCTGTGACTCTGCACATGTACTCTGCAATTCGA 368
|||||
Db 121 TGTTCATCAACGTGCTGAGCAGTCCGCTGTGACTCTGCACATGTACTCTGCAATTCGA 180
QY 369 TGGAGTCTCTTACCTCCGATCATGAGACCTTACGAGCGAGTACGGGTGTGAGACCC 428
|||||
Db 181 TGGAGTCTCTTACCTCCGATCATGAGACCTTACGAGCGAGTACGGGTGTGAGACCC 240
QY 429 TGAATATGCTGCTACCCAGCTAGCTCAACCAACCATAGATGAGATGAGCTCGGCAAACTCTC 488
|||||
Db 241 TGAATATGCTGCTACCCAGCTAGCTCAACCAACCATAGATGAGATGAGCTCGGCAAACTCTC 300
QY 489 TTTGGACAAAGTCTTCCGGGAACGGAGTCCGATTCGACAGATTGTGATCCCATCAA 548
|||||
Db 301 TGTGGACAAAGTCTTCCGGGAACGGAGTCCGATTCGACAGATTGTGATCCCATCAA 360
QY 549 CCAAGCTGCTGACTGCTGGGGTATCCGCTGCCCTNGCTTATGAGATCAAGGATATCCATGT 608
|||||
Db 361 CCAAGCTGCTGACTGCTGGGGTATCCGCTGCCCTNGCTTATGAGATCAAGGATATCCATGT 420
QY 609 GCCACCCCGGGTGAAGAGTCTATGCGAGATGCGAGTGAAGCGAGCGCGGGAACGGGC 668
|||||
Db 421 GCCACCCCGGGTGAAGAGTCTATGCGAGATGCGAGTGAAGCGAGCGCGGGAACGGGC 480
QY 669 CACAGTTCTAGAGTCTGAGGGGACCCGAGAGTCCGCTCATATGTGCGAGAGGAGAA 728
|||||
Db 481 CACAGTTCTAGAGTCTGAGGGGACCCGAGAGTCCGCTCATATGTGCGAGAGGAGAA 540
QY 729 ACAGGCCCAAGTCTGCTGCCCTCCGAGCAGAAAAGGCTGAACAGATTAATCAAGCAGCAGG 788
|||||
Db 541 ACAGGCCCAAGTCTGCTGCCCTCCGAGCAGAAAAGGCTGAACAGATTAATCAAGCAGCAGG 600
QY 789 AGAGGCCCAAGTCTGCTGCCCTCCGAGCAGAAAAGGCTGAAGCTGAATTCGATTCCTGGC 848
|||||
Db 601 AGAGGCCCAAGTCTGCTGCCCTCCGAGCAGAAAAGGCTGAAGCTGAATTCGATTCCTGGC 660
QY 849 TGCAGCTCTGACACACATTAATGAGATGAGAGTCTCACTACGTGTGCTCAGACAGTA 908
|||||
Db 661 TGCAGCTCTGACACACATTAATGAGATGAGAGTCTCACTACGTGTGCTCAGACAGTA 720
QY 909 TGTACAGCGGCTTCTCAAACTGGCCAAAGGCTCAACACATCTCTACTGGCTCTCAACC 968
|||||
Db 721 TGTACAGCGGCTTCTCAAACTGGCCAAAGGCTCAACACATCTCTACTGGCTCTCAACC 780
QY 969 TGGCGATGTTCACAGCATGGTGGCTCAGGCGCATGGGTATATGAGACCCCTCACCAAGC 1028
|||||
Db 781 TGGCGATGTTCACAGCATGGTGGCTCAGGCGCATGGGTATATGAGACCCCTCACCAAGC 840
QY 1029 CCCAGTCCAGGAGTCTCAGACACTCTCTCCATGGGAGAGAGAGATGTCCAGGGTAC 1088
|||||
Db 841 CCCAGTCCAGGAGTCTCAGACACTCTCTCCATGGGAGAGAGAGATGTCCAGGGTAC 900
QY 1089 AGATGCAGTNTGTAGAGGAGTGTGAGTCAAA 1124
|||||
Db 901 AGATGCAGTNTGTAGAGG-ACCTGTGATGAGTTAA 935

RESULT 3
BM476304 1110 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT 6479118 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5558923
DEFINITION 5' mRNA sequence.
ACCESSION BM476304
VERSION BM476304.1 GI:18525346
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE 1 (bases 1 to 1110)
AUTHORS NIH-MGC http://imgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://imgc.lnl.gov
Plate: LLM12282 row: k column: 20
High quality sequence start: 8
High quality sequence stop: 678.
Location/Qualifiers
1. 1110
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5558923"
/clone_lib="NIH_MGC_88"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: small intestine; Vector: PCMV-SPOB6;
Site 1: NotI; Site 2: SalI; Cloned unidirectionally;
oligo-dT primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life
technologies. Note: this is a NIH_MGC Library."
BASE COUNT 249 a 322 c 213 t
ORIGIN

Query Match 75.7% Score 899.4; DB 10; Length 1110;
Best Local Similarity 95.0%; Pred. No. 1,4e-209;
Matches 993; Conservative 0; Mismatches 43; Indels 9; Gaps 6;

QY 36 GTTCCGAGAGTCTGCTGGGGGCGTGGGAAATGCTGGCGCGCGGGGCGGAGCATGGG 95
|||
Db 27 GTCCGGAGAGTCTGCTGGGGGCGTGGGAAATGCTGGCGCGCGGGGCGGAGCATGGG 85
QY 96 CCTTTTGTGAGAGGGCTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 154
|||
Db 86 CCTTTTGTGAGAGGGCTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 145
QY 155 GATTGCCCCGAAACACCGTGTACTGTTGCTGCGCGAGGAGAGGCTGGTGGTGGAGC 214
|||||
Db 146 GATTGCCCCGAAACACCGTGTACTGTTGCTGCGCGAGGAGAGGCTGGTGGTGGAGC 205
QY 215 GAATGGGCCGATTTCACACCGGATCCGAGGCTGGTGGTGAACATCTCATCTGTGTAG 274
|||||
Db 206 GAATGGGCCGATTTCACACCGGATCCGAGGCTGGTGGTGAACATCTCATCTGTGTAG 265
QY 275 ACCGAGATCGATATGTGACAGAGTCTCAGGAATTTGTCATCAACGTGAGCAGTGG 334
|||||
Db 266 ACCGAGATCGATATGTGACAGAGTCTCAGGAATTTGTCATCAACGTGAGCAGTGG 325
QY 335 CTGAGACTCTGCAAAATGTAAGTCTGCAAAATGTAAGTCTTACTGCGGATCATGG 394
|||||
Db 326 CTGAGACTCTGCAAAATGTAAGTCTGCAAAATGTAAGTCTTACTGCGGATCATGG 385
QY 395 ACCCTTACAAGGCAAGCTAGCGTGTGAGAGACCTGTGAGTATGCGGTACCCAGCTAGCTC 454
|||||
Db 386 ACCCTTACAAGGCAAGCTAGCGTGTGAGAGACCTGTGAGTATGCGGTACCCAGCTAGCTC 445
QY 455 AAACACCATGAGATGAGAGTCTGCGCAAACTCTCTTNTGACAAAGTCTTCCGGGAACGGG 514
|||||
Db 446 AAACACCATGAGATGAGAGTCTGCGCAAACTCTCTTNTGACAAAGTCTTCCGGGAACGGG 505
QY 515 AGTCCCTGAATGCCAGATGAGTGTGATGATCAACCAAGCTGCTGATGCTGGGGGTATCC 574
|||||
Db 506 AGTCCCTGAATGCCAGATGAGTGTGATGATCAACCAAGCTGCTGATGCTGGGGGTATCC 565
```

575 GCTGCTTCGTTATGATGATCAAGATATCCATGTCGCCACCCCGGGTGAAAGATCTATGC 634
|||||
566 GCTGCTCCGTTATGATGATCAAGATATCCATGTCGCCACCCCGGGTGAAAGATCTATGC 625
|||||
635 AGATCAGGTGAGAGCAGACGCCGGGAAAGCGGCCACAGTTCTAGAGTGTGAGGGAGCC 694
|||||
636 AGATCAGGTGAGAGCAGACGCCGGGAAAGCGGCCACAGTTCTAGAGTGTGAGGGAGCC 685
|||||
695 GAGATCGGCCCTCATATGTCGGCAGAGGAAGAGAGCCAGATCTGCGCTCCGAG 754
|||||
666 GAGATCGGCCCTCATATGTCGGCAGAGGAAGAGAGCCAGATCTGCGCTCCGAG 745
|||||
755 CAGAAAAGGCTGACAGATAAATCAGGAGAGAGAGCCAGTGCAGTTCTGCGCAG 814
|||||
746 CAGAAAAGGCTGACAGATAAATCAGGAGAGAGAGCCAGTGCAGTTCTGCGCAG 805
|||||
815 CCAAGGCTTAAGCTGAGTATTCGAATCTGCTGCGAGCTTGACACACATTAATGAG 874
|||||
806 CCAAGGCTTAAGCTGAGTATTCGAATCTGCTGCGAGCTTGACACACATTAATGAG 865
|||||
875 ATGACAGAGCTTCACTGACTGTGGCGAGACAGTATGTACAGCGG-TTCTCCAAAGTGGC 933
|||||
866 ATGCGCGAGCTTCCCTGACTGTGGCGAGACAGTATGTACAGCGGTTCTCCAAAGTGGC 925
|||||
934 AAGAGCTCCACACTATCTACTGCTCCCAACCTGCGG-ATGTCACAGCAGATGCTG 992
|||||
926 AAGAGCTCCACACTATCTACTGCTCCCAACCTGCGGAGAGCGCCAGCATGAGGAGC 985
|||||
993 TCA--GGCATGCGGTATATGAGAGCCCTCACC---AAGCCCCAGTGGCAGAGACTGCA 1047
|||||
986 TCAAGGCTCCGCGGTATGAGAACCCCTCCCAAAAGCCCCAGGCGCAGAGACTGCC 1045
|||||
1048 GACTCCTCTCCAGTGTGAGAGCAGCA 1072
|||||
1046 CGACTCCCTCCCGGGGAGCA 1070
|||||

RESULT 4
AL516503/c 962 bp mRNA linear EST 13-FEB-2001
LOCUS AL516503 LTI_NFL011.NBC1 Homo sapiens cDNA clone CS0DA006YH04 3
DEFINITION
AL516503
prime, mRNA sequence.
ACCESSION
AL516503
VERSION
AL516503.1 GI:12779996
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 962)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished (2001)
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1. 962
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DA006YH04"
/clone_11b="LTI_NFL011.NBC1"
/sex="male"
/tissue.type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-Oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library is not normalized, but is the control for
the normalized libraries. Library was constructed by Life
Technologies. Contact : Feng Liang Life Technologies, a

division of Invitrogen 9800 Medical Center Drive Rockville
, Maryland 20850, USA Fax : (1) 301 610 8371 Email :
fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 193 a 262 c 248 g 243 t 16 others
ORIGIN

Query Match 74.4% Score 884.2; DB 9: Length 962;
Best Local Similarity 95.8%; Pred No. 6.8e-206;
Matches 920; Conservative 13; Mismatches 23; Indels 4; Gaps 3;

232 CGATCTCTGAGCCCTGGTTGAACATCTATCCCTGTGTAGACCGGATCCGATATGTG 291
|-:|||||
962 CGATCTCTGAGCCCTGG-TTGGACATCTCTATCCCTGTGTAGACCGGATCCGATATGTG 904
|-:|||||
292 CAGAGTCTAAGGAATTTGTCAATCAAGTCTGCTGAGCAAGTCCGCTGTGACTCCGCAAT 351
|||:|||||
903 CAGACTCTCAAGGAATTTGTCAATCAAGTCTGCTGAGCAAGTCCGCTGTGACTCCGCAAT 844
|||:|||||
352 GTAACCTGCAATATGATGAGAGGCTTTTACCTGCGCATCATGAGACCTTACAAAGGCAAGC 411
|||:|||||
843 GTAACCTGCAATATGATGAGAGGCTTTTACCTGCGCATCATGAGACCTTACAAAGGCAAGC 784
|||:|||||
412 TACGGTGTGAGAGACCTGATGATGCGCTACACCACTAGCTCAAAACCATGAGATCA 471
|||:|||||
783 TACGGTGTGAGAGACCTGATGATGCGCTACACCACTAGCTCAAAACCATGAGATCA 724
|||:|||||
472 GAGCTCGGCAAACTCTCTTTGAGCAAAAGTCTCCGGAAACGGAGTCCCTGAATGCGAC 531
|||||
723 GAGCTCGGCAAACTCTCTTTGAGCAAAAGTCTCCGGAAACGGAGTCCCTGAATGCGAC 664
|||||
532 ATTGTGATGCCATCAACCAAGCTGCTGAGTCTGGGGATATCCGCTGGCTTGGTTATGAG 591
|||||
663 ATTGTGATGCCATCAACCAAGCTGCTGAGTCTGGGGATATCCGCTGGCTTGGTTATGAG 604
|||||
592 ATCAAGATATCATGATGTCACCCCGGGTGAAGATCTATGCAGATGAGTGAAGCA 651
|||||
603 ATCAAGATATCATGATGTCACCCCGGGTGAAGATCTATGCAGATGAGTGAAGCA 544
|||||
652 GAGCGCGGAAACGGGCCACAGTTCTAGAGTGTGAGGGAGCCGAGAGTCCGCTATCAT 711
|||||
543 GAGCGCGGAAACGGGCCACAGTTCTAGAGTGTGAGGGAGCCGAGAGTCCGCTATCAT 484
|||||
712 GTGGCAGAGGGAAGAAAGAGCCCGATCTGCGCTCCGAGCAAGAAAGGCTGAAGCA 771
|||||
483 GTGGCAGAGGGAAGAAAGAGCCCGATCTGCGCTCCGAGCAAGAAAGGCTGAAGCA 424
|||||
772 ATPAATCAGGACAGAGAGAGGAGGCTGAG-TTCTGGCGAAAGGCCAAGGCTPAAGCTGA 830
|||||
423 ATPAATCAGGACAGAGAGAGGAGGCTGAG-TTCTGGCGAAAGGCCAAGGCTPAAGCTGA 364
|||||
831 AGCTATTCGAATCTGCGCTGAGCTGTGACACAACATTAATGAGAGTGCAGAGCTTCACT 890
|||||
363 AGCTATTCGAATCTGCGCTGAGCTGTGACACAACATTAATGAGAGTGCAGAGCTTCACT 304
|||||
891 GAGTGTGGCGAGAGTATGTCAGCGCTTCTCAAACTGCGCAAGAGTCCAAACTAT 950
|||||
303 GAGTGTGGCGAGAGTATGTCAGCGCTTCTCAAACTGCGCAAGAGTCCAAACTAT 244
|||||
951 CCTACTGCGCTCCAAACCTGCGGATGTCCACAGCATGTGCTCAGGCGCATGGGTATATA 1010
|||:|||||
243 CCAACTGCGCTCCAAACCTGCGGATGTCCACAGCATGTGCTCAGGCGCATGGGTATATA 184
|||:|||||
1011 TGGAGCCCTCACCAGAGCCCGGAGGAGTCCAGAGCTCAGAGTCTACTCTCCAGTGGAGCAG 1070
|||||
183 TGGAGCCCTCACCAGAGCCCGGAGGAGTCCAGAGCTCAGAGTCTACTCTCCAGTGGAGCAG 124
|||||
1071 CAGAGATGTCAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1130
|||||
123 CAGAGATGTCAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 64
|||||
1131 TTAGTGAAGTGGGCTTNGCCAGGAGTGTGGGAGCAAGG--AAGCAGATTTTCTGATTT 1188
|||||

```
DB 63 TTAGTGAGCTGGGCTTGCCAGGAGATYGGGACAAAGATATGCAGATTTCCTGTT 4
|||||
RESULT 5
BM468527 994 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6475743 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5578198
DEFINITION 5', mRNA sequence.
ACCESSION BM468527
VERSION BM468527.1 GI:18517569
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 994)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DPF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12332 row: n column: 23
High quality sequence stop: 658.
Location/Qualifiers
1..994
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5578198"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 225 a 276 c 292 g 194 t 7 others
ORIGIN
Query Match 73.4%; Score 871.8; DB 10; Length 994;
Best local Similarity 97.5%; Pred. No. 7.6e-203;
Matches 912; Conservative 0; Mismatches 20; Indels 3; Gaps 3;

QY 49 CTGCGGCGGGGGAATGCTGGCGGGCGGGGCGGACACGAGGCGCTTTTGTGAG 108
|||||
DB 1 CTGCGGCGGGGGAATGCTGGCGGGCGGGCGGCGGCGG-GGGSGCATGGGCGCTTTTGTGAG 59
|||||
QY 109 GGGCTCTACTGCTTCTTGGCGCGCTCCG-GCGCGGCTCTCTGATGGCCGAAA 167
|||||
DB 60 GGGCTCTACTGCTTCTTGGCGCGCTCCGCGCGGCGCTCTCTGATGGCCGAAA 119
|||||
QY 168 CACCGTGTACTGCTTCTTGGCGCGGAGGCGCTGGTGTGAGCGAAATGGGCGGATT 227
|||||
DB 120 CACCGTGTACTGCTTCTTGGCGCGGAGGCGCTGGTGTGAGCGAAATGGGCGGATT 179
|||||
QY 228 CACCGGATCTGGAGCGCTGTTGAAATCCATCCCTGTTTAAACCGGATCCGATA 287
|||||
DB 180 CACCGGATCTGGAGCGCTGTTGAAATCCATCCCTGTTTAAACCGGATCCGATA 239
|||||
QY 288 TGTGAGAGTCTCAAGAAATTTGATCAACGTGCTGAGCAGTGGCTGTGACTTCGA 347
|||||
DB 240 TGTGAGAGTCTCAAGAAATTTGATCAACGTGCTGAGCAGTGGCTGTGACTTCGA 299
|||||
QY 348 CAATGTACTCTGCAATTCATGAGTCTTTTACTCTGCGATCATGACCTTACAGGC 407
|||||
```

```
DB 300 CAATGTACTCTGCAATTCATGAGTCTTTTACTCTGCGATCATGAGACCTTACAGGC 359
|||||
QY 408 AAGCTACGCTGTGAGAGACCTTGAGTATGCCGTACACCAGTACTGTAACAAACCATGAG 467
|||||
DB 360 AAGCTACGCTGTGAGAGACCTTGAGTATGCCGTACACCAGTACTGTAACAAACCATGAG 419
|||||
QY 468 ATCAGAGCTGGCAAACTCTCTNTGACAAAGATCTTCGGGAAAGGAGTCCCTGAATGC 527
|||||
DB 420 ATCAGAGCTGGCAAACTCTCTNTGACAAAGATCTTCGGGAAAGGAGTCCCTGAATGC 479
|||||
QY 528 CAGCATTTGTGAGTCCATCAACCAAGCTGCTGACTGTGGGGTATCCGTGCTTCGTTA 587
|||||
DB 480 CAGCATTTGTGAGTCCATCAACCAAGCTGCTGACTGTGGGGTATCCGTGCTTCGTTA 539
|||||
QY 588 TGAGATCAAGGATATCCATGTGGCCACCCGGGTAAAGAGTCTATGAGATGCAAGTGGGA 647
|||||
DB 540 TGAGATCAAGGATATCCATGTGGCCACCCGGGTAAAGAGTCTATGAGATGCAAGTGGGA 599
|||||
QY 648 GCGAGACGGCGGAAACGGGCCACAGTTCTAGAGTCTGAGGGGACCCGAGTGGCCAT 707
|||||
DB 600 GCGAGACGGCGGAAACGGGCCACAGTTCTAGAGTCTGAGGGGACCCGAGTGGCCAT 659
|||||
QY 708 CAATGTGGCAGAAAGGAAACAGGCCACAGATCTTGCGCTCCGAGCAGAAAGCTGA 767
|||||
DB 660 CAATGTGGCAGAAAGGAAACAGGCCACAGATCTTGCGCTCCGAGCAGAAAGGCTGA 719
|||||
QY 768 ACAGATTAATCAGGACAGAGAGGCGGACGTCTGCGGGAAGGCCAAGGCTTAAGC 827
|||||
DB 720 ACAGATTAATCAGGACAGAGAGGCGGACGTCTGCGGGAAGGCCAAGGCTTAAGC 779
|||||
QY 828 TGAAGCTATTCGAATCCTGCTGAGCTGTGACACACATTAATGAGATGAGCAGGCTTC 887
|||||
DB 780 TGAAGCTATTCGAATCCTGCTGAGCTGTGACACACATTAATGAGATGAGCAGGCTTC 839
|||||
QY 888 ACTGACTGTGGCGGAGCAGATATGTCTACGCGGTTCTCCAACTGGCCCAAGACTCAACAC 947
|||||
DB 840 ACTGACTGTGGCGGAGCAGATATGTCTACGCGGTTCTCCAACTGGCCCAAGACTCAACAC 899
|||||
QY 948 TATCTTACTG-CCCTCCACCCCTGGGAGTGCACC 981
|||||
DB 900 TATCTTACTGCCCCCTCAAAACCCCTGGAGATGTCAAC 934
|||||

RESULT 6
BM468407 998 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6475436 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5578133
DEFINITION 5', mRNA sequence.
ACCESSION BM468407
VERSION BM468407.1 GI:18517449
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 998)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DPF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12332 row: 1 column: 06
High quality sequence stop: 656.
Location/Qualifiers
1..998
/organism="Homo sapiens"
SOURCE
```


ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1011)
NIH-MGC <http://mgc.ncl.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC

FEATURES

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLM12137 row: n column: 16
High quality sequence stop: 622.
Location/Qualifiers
1..1011

BASE COUNT 232 a 289 c 288 g 201 t 1 others
ORIGIN
Query Match 70.3% Score 835.4; DB 10; Length 1011;
Best Local Similarity 98.1% Pred. No. 6.3e-194;
Matches 875; Conservative 0; Mismatches 14; Indels 3; Gaps 3;

84 GGGGACTGGGCGCCCTTTGCTGAGGGGCTCTACTGCTTCTGCGCGCTCG-GCC 142
|||||
4 GGGGACTGGGCGCCCTTTGCTGAGGGGCTCTACTGCTTCTGCGCGCTCGCGCC 63
|||||
143 GGGCTCTCTGATTTGGCGGAAACCGGTGATCTTCTGCGCGCGAGGAGGCT 202
|||||
64 GGGCTCTCTGATTTGGCGGAAACCGGTGATCTTCTGCGCGCGAGGAGGCT 123
|||||
203 GGGTGTGAGGAAATGGGCGCATTCACCGATCTGAGGCTGTTGAACATCTCA 262
|||||
124 GGGTGTGAGGAAATGGGCGCATTCACCGATCTGAGGCTGTTGAACATCTCA 183
|||||
263 TCCCTGTGTAGACCGGATCCGATATGTCAGAGTCTCAAGCAAAATGTCATCACGTGC 322
|||||
184 TCCCTGTGTAGACCGGATCCGATATGTCAGAGTCTCAAGCAAAATGTCATCACGTGC 243
|||||
323 CTGAGAGTGGGCTGTGATCTCGAATGTAATCTGCAAAATGAGAGTCTTACC 382
|||||
244 CTGAGAGTGGGCTGTGATCTCGAATGTAATCTGCAAAATGAGAGTCTTACC 303
|||||
383 TCGCATCTATGACCTTCAAGCAAGCTAGCGTGTGAGAGCGCGAGTATGCCGTCA 442
|||||
304 TCGCATCTATGACCTTCAAGCAAGCTAGCGTGTGAGAGCGCGAGTATGCCGTCA 363
|||||
443 CCCAGCTAGCTCAAAACACATGATGATCGAGAGTGGCAAACTCTGTTGAGCAAAAGTCT 502
|||||
364 CCCAGCTAGCTCAAAACACATGATGATCGAGAGTGGCAAACTCTGTTGAGCAAAAGTCT 423
|||||
503 TCGGGAAGGAGTCCCTGATGATCCAGCATTTGTGATGCAATCAAGCTGCTGACT 562
|||||
424 TCGGGAAGGAGTCCCTGATGATCCAGCATTTGTGATGCAATCAAGCTGCTGACT 483
|||||
563 GCGGGGATTCGCGCTGCTTCGTTATGAGATCAAGGATATCCATGAGCCGCCCGGCTGA 622
|||||
484 GCTGGGGATTCGCGCTGCTTCGTTATGAGATCAAGGATATCCATGAGCCGCCCGGCTGA 543

623 AAGACTATATGATGATGAGTGGAGGCGAGCGGCGGAAACGGGCCACAGTTTAGACT 682
|||||
544 AAGACTATATGATGATGAGTGGAGGCGAGCGGCGGAAACGGGCCACAGTTTAGACT 603
|||||
683 CTGAGGGGAGCCGAGAGTGTGGCATCAATGTGGCAGAGGAGGAAACGAGCCGAGTCC 742
|||||
604 CTGAGGGGAGCCGAGAGTGTGGCATCAATGTGGCAGAGGAGGAAACGAGCCGAGTCC 663
|||||
743 TGGCTCCGAGACAGAAAGGCTGAACAGATTAATCAGCAGCAGAGAGGAGCCAGTGCAG 802
|||||
664 TGGCTCCGAGACAGAAAGGCTGAACAGATTAATCAGCAGCAGAGAGGAGCCAGTGCAG 723
|||||
803 TTCTGGCGAAGGCCAAGCTAAAGCTGAATTCCTGCTGCTGACCTGACAC 862
|||||
724 TTCTGGCGAAGGCCAAGCTAAAGCTGAATTCCTGCTGCTGACCTGACAC 783
|||||
863 AACATTAATGAGATGAGAGCTTACAGCTGAGGCGGAGAGTATGACAGCGGCTTCT 922
|||||
784 -ACATATTGAGATGAGAGCTTACAGCTGAGGCGGAGAGTATGACAGCGGCTTCT 842
|||||
923 CCAACTGGCGCAAGAGCT-CCAACACTATCTACTGCGCTCCCAACCCCTGGCG 973
|||||
843 CCCAAGTGGCGCAAGAGCTCCCAACACTATCTTCTGCGCTCCCAACCCCTGGCG 894
|||||

RESULT 9
AL525168 849 bp mRNA linear EST 13-FEB-2001
LOCUS AL525168 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC005YF20 5
DEFINITION prime, mRNA sequence.
ACCESSION AL525168
VERSION AL525168.1 GI:12788661
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 849)
AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segregenoscope.cns.fr; Web : www.genoscope.cns.fr.

FEATURES

source

1..849
Location/Qualifiers
1..849
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DC005YF20"
/clone_lib="LTI_NFL003_NBC3"
/sex="male"
/lissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with NotI and
cloned into the NotI and EcoRV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact: Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax: (1) 301 610
8371 Email: liang@lifetech.com
<http://fulllength.invitrogen.com>"

BASE COUNT 195 a 228 c 260 g 165 t 1 others
ORIGIN
Query Match 69.1% Score 821.4; DB 9; Length 849;
Best Local Similarity 99.3% Pred. No. 1.5e-190;
Matches 844; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

53 GCGGTGGGAATGCTGGCGCGCGCGGGGCGACTGTGGGCGCTTTTGTGAGGGC 112
|||||
1 GCGGTGGGAATGCTGGCGCGCGCGGGGCG-GAGCGCATGGGGCCCTTTGTGAGGGC 59
|||
113 TGTCTACTGGCTTCTGCGCGCGCGCTCCG-GCGCGCTCTCTGTGATGGCCGGAACAC 171
|||||
60 TGTCTACTGGCTTCTGCGCGCGCGCTCCGCGCGCTCTCTGTGATGGCCGGAACAC 119
|||||
172 GTGACTCTGTGCGCGCGCGAGAGAGGCGGTGGTGGAGGAATGGCGCATTCAC 231
|||||
120 GTGACTCTGTGCGCGCGAGAGAGGCGGTGGTGGAGGAATGGCGCATTCAC 179
|||||
232 CGGATCTGAGACCTGTGTTGAACATCTCATCTCTGTGTTAGACCGGATCCGATATGTG 291
|||||
180 CGGATCTGAGACCTGTGTTGAACATCTCATCTCTGTGTTAGACCGGATCCGATATGTG 239
|||||
292 CAGATCTCAAGGAATTTTCATCAACGCTGCTGAGCACTCGGCTGTGACTCTCCGCAAT 351
|||||
240 CAGATCTCAAGGAATTTTCATCAACGCTGAGCACTCGGCTGTGACTCTCCGCAAT 299
|||||
352 GTAACTCTCAAGGAATTTTCATCAACGCTGAGCACTCGGCTGTGACTCTCCGCAAGC 411
|||||
300 GTAACTCTCAAGGAATTTTCATCAACGCTGAGCACTCGGCTGTGACTCTCCGCAAGC 359
|||||
412 TACGGTGTGAGAGACCTGTGATGCGCTCACCAAGCTCAACCAACCAATGATGATCA 471
|||||
360 TACGGTGTGAGAGACCTGTGATGCGCTCACCAAGCTCAACCAACCAATGATGATCA 419
|||||
472 GAGCTCGGCAAACTCTCTTTGGAACAAGTCTTCGGGAAACGGGAGTCCCTGAAATGCCAC 531
|||||
420 GAGCTCGGCAAACTCTCTTTGGAACAAGTCTTCGGGAAACGGGAGTCCCTGAAATGCCAC 479
|||||
532 ATTGTGATGATGATCAACCAAGCTGTGATGCGGTGATGCGGTGATGCGGTGATGAG 591
|||||
480 ATTGTGATGATGATCAACCAAGCTGTGATGCGGTGATGCGGTGATGCGGTGATGAG 539
|||||
592 ATCAAGGATATCATGATGCGCACCCCGGATGAAGATCATGATGATGATGATGAGGA 651
|||||
540 ATCAAGGATATCATGATGCGCACCCCGGATGAAGATCATGATGATGATGATGAGGA 599
|||||
652 GAGCGGGGAAACGGGCGCACATTTCTAGAGTCTGAGGGAGCCGAGAGTCCGCAATCAAT 711
|||||
600 GAGCGGGGAAACGGGCGCACATTTCTAGAGTCTGAGGGAGCCGAGAGTCCGCAATCAAT 659
|||||
712 GTGGCAGAGGGAAGAAACAGGCGCATCTGGCTCCGAAAGCAGAAAAGCTGAACAG 711
|||||
660 GTGGCAGAGGGAAGAAACAGGCGCATCTGGCTCCGAAAGCAGAAAAGCTGAACAG 719
|||||
772 ATAAATCAGAGCAGAGAGAGGCGCATGATGATGATGATGATGATGATGATGATGAT 831
|||||
720 ATAAATCAGAGCAGAGAGAGGCGCATGATGATGATGATGATGATGATGATGATGAT 779
|||||
832 GCTATTCGAATCTGCTGCTGAGCTCTGACACAAATATGAGATGAGATGAGATGAGATGAG 891
|||||
780 GCTATTCGAATCTGCTGCTGAGCTCTGACACAAATATGAGATGAGATGAGATGAGATGAG 839
|||||
892 ACTGTGGCG 901
|||||
840 ACTGTGGCG 849

RESULT 10

BM453159 1061 bp mRNA linear EST 05-FEB-2002
LOCUS BM453159
DEFINITION AGENCOURT_6387825 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5529785
5', mRNA sequence.
ACCESSION BM453159
VERSION BM453159.1 GI:18502199
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1061)
AUTHORS NIH-MGC <http://mgi.nhl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LAM12208 row: m column: 18
High quality sequence stop: 661.
Location/Qualifiers
1..1061
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5529785"
/clone.lib="NIH_MGC_71"
/tissue_type="leiomyosarcoma"
/lab_nos="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."
BASE COUNT 259 a 298 c 279 g 225 t
ORIGIN
Query Match 69.1%; Score 821; DB 10; Length 1061;
Best local Similarity 98.8%; Pred. No. 2..2e-190;
Matches 835; Conservative 0; Mismatches 9; Indels 1; Gaps 1;
1: 1:
QY 345 CGCAATGTAATCTGCAATGATGAGTCTTTTACCTGCGCATGAGACCTTACAA 404
|||||
DB 1 CGCAATGTAATCTGCAATGATGAGTCTTTTACCTGCGCATGAGACCTTACAA 60
|||||
QY 405 GGCACCTAGGCTGTGAGAGACCTGATGATGCGCTGACCTGATGATGATGATGATGAT 464
|||||
DB 61 GGCACCTAGGCTGTGAGAGACCTGATGATGCGCTGACCTGATGATGATGATGATGAT 120
|||||
QY 465 GAGATGAGAGCTGCGCAACTCTTNTGACAAGTCTTCCGGGAAACGGGAGTCCCTGAA 524
|||||
DB 121 GAGATGAGAGCTGCGCAACTCTCTCGACAAGTCTTCCGGGAAACGGGAGTCCCTGAA 180
|||||
QY 525 TGCACAGATTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 584
|||||
DB 181 TGCACAGATTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
|||||
QY 585 TTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 644
|||||
DB 241 TTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
|||||
QY 645 GGAGGAGAGAGCGGCGGAAACGGGCGACAGTTCTAGAGTCTGAGGGAGCCGAGAGTCCGGC 704
|||||
DB 301 GGAGGAGAGAGCGGCGGAAACGGGCGACAGTTCTAGAGTCTGAGGGAGCCGAGAGTCCGGC 360
|||||
QY 705 CATCAATGTGTGCGAAGGGAAGAAACAGGCGCGCATGCTGCGCTCCGAGACAAAAAGGC 784
|||||
DB 361 CATCAATGTGTGCGAAGGGAAGAAACAGGCGCGCATGCTGCGCTCCGAGACAAAAAGGC 420
|||||
QY 765 TGAACAGATTAATCAGGACAGAGAGGCGCATGCTGCGGAGGCGCAAGGCTAA 824
|||||
DB 421 TGAACAGATTAATCAGGACAGAGAGGCGCATGCTGCGGAGGCGCAAGGCTAA 480
|||||
QY 825 AGCTGAAGTATTCGATCTGCTGAGCTCTGACACAAATATGAGATGAGATGAGATGAGATGAG 884
|||||
DB 481 AGCTGAAGTATTCGATCTGCTGAGCTCTGACACAAATATGAGATGAGATGAGATGAGATGAG 540
|||||
QY 885 TTTCACGTAGCTGTGGCGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 944
|||||
DB 541 TTTCACGTAGCTGTGGCGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
|||||

OY	239	TGGAGCTGCGTTTGAACAATCCATCCATCCCTGTGTATTAGACCGGATCCGATATGTGCAGAATC	298
Db	73	TGGACCCTGGTTTTGAACATCCATCCATCCCTGTGTATTAGACCGGATCCGATATGTGCAGAATC	132
OY	299	TCAAGGAAATTTGTCAATCAAGGCGCTTGAGAGAGTCGGCTGTGACTCTGCACAAATGTAATC	358
Db	133	TCAAGGAAATTTGTCAATCAAGGCGCTTGAGAGAGTCGGCTGTGACTCTGCACAAATGTAATC	192
OY	359	TGCAATCATGATGAGTAGCTCTTTACCTGCGCATCATGAGACCCCTTACAAAGGCAAGCTAACGGTG	418
Db	193	TGCAATCATGATGAGTAGCTCTTTACCTGCGCATCATGAGACCCCTTACAAAGGCAAGCTAACGGTG	252
OY	419	TGGAGGACCCTGAGTATGCGCGTCACCCAGCTAAGCTCAACAACCACTTGATATCAGAGCTCG	478
Db	253	TGGAGGACCCTGAGTATGCGCGTCACCCAGCTAAGCTCAACAACCACTTGATATCAGAGCTCG	312
OY	479	GCAAACTCTCTTTGAGACAAGATCTTCGGGAAAGGGGAGTCCCTGAAATGCCAGCATATGGG	538
Db	313	GCAAACTCTCTTTGAGACAAGATCTTCGGGAAAGGGGAGTCCCTGAAATGCCAGCATATGGG	372
OY	539	ATGCCATCAACCAAGCTGCTGACTGCTGCGGTATCCGCTGCTTNCGTTATGATGATCAAG	598
Db	373	ATGCCATCAACCAAGCTGCTGACTGCTGCGGTATCCGCTGCTTATGATGATCAAG	432
OY	599	ATATCATGTGCTCCACCCCGGGGTGAAGAAGTCTATGCAATGCAAGTGGAGGACAGCGGC	658
Db	433	ATATCATGTGCTCCACCCCGGGGTGAAGAAGTCTATGCAATGCAAGTGGAGGACAGCGGC	492
OY	659	GGAAGCGGGCCACAGTTCTAGAGTCTGAGGGGAGCCGAGAGTCCGCGCATATGTGGAG	718
Db	493	GGAAGCGGGCCACAGTTCTAGAGTCTGAGGGGAGCCGAGAGTCCGCGCATATGTGGAG	552
OY	719	AAGGAGAGAAACAGGCCCCAGATCTGTGGCTCCGGAAGCAGAAAAGGCTGAACAGATAATC	778
Db	553	AAGGAGAGAAACAGGCCCCAGATCTGTGGCTCCGGAAGCAGAAAAGGCTGAACAGATAATC	612
OY	779	AGCGACGACGAGAGGCGCACAGTGTCTGGGAGAGCCCAAAGCTTAAAGCTGAAGCTATTTC	838
Db	613	AGCGACGACGAGAGGCGCACAGTGTCTGGGAGAGCCCAAAGCTTAAAGCTGAAGCTATTTC	672
OY	839	GAATCTGGCTCTCAGCTCTGACAAACATTAATGAGATGACGACACTTCACTGACTGTGG	898
Db	673	GAATCTGGCTCTCAGCTCTGACAAACATTAATGAGATGACGACACTTCACTGACTGTGG	732
OY	899	CCGAGCAGTATGTACAGCGCGTCTCCAAACTGCGCAAGGACTCCAAACTATCTACTATGC	958
Db	733	CCGAGCAGTATGTACAGCGCGTCTCCAAACTGCGCAAGGACTCCAAACTATCTACTATGC	792
OY	959	CCTCCACACCTTGG--CGATGTCAACAGCATGTTGGCTCAG---CCATGGGTGTATATGGA	1014
Db	793	CCTCCACACCTTGGCGGATGTCAACAGCATGTTGGGCTTAAAGCCCATGGGTGTATATGGA	852
OY	1015	---GGCCTTACCAAAAGCCCC	1031
Db	853	AGCCCCCTTACCAAAAGCCCC	872
RESULT 14			
LOCUS	BG479080	943 bp	mRNA linear EST 21-MAR-2001
DEFINITION	602326159t1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:4649846 5'		
ACCESSION	BG479080		
VERSION	BG479080.1	GI:13411359	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE	1 (bases 1 to 943)		
AUTHORS	NIH-MGC http://mgc.mcl.nih.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		

COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
plate: L1CML1430 row: e column: 15
High quality sequence stop: 857.

FEATURES
Source
Location/Qualifiers
1. 943
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4649846"
/clone_lib="NIH.MGC.21"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: placenta; Vector: pOTB7; Site:1: XhoI;
Site:2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the Laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 210 a 254 c 290 g 189 t

ORIGIN

Query Match 66.6%; Score 791.4; DB 10; Length 943;
Best Local Similarity 98.4%; Pred. No. 3.6e-183;
Matches 861; Conservative 0; Mismatches 8; Indels 6; Gaps 5;

32 GTTGGTTCGGAGGTCGCTGCGCGCGGTGGGAATCTGGCGCGCGGGGGGCACT 91
|||||
2 GTTGGTTCGGAGGTCGCTGCGCGCGGTGGGAATCTGGCGCGCGGGGGCACT 60
92 GGGGCCCTTTTCTAGAGGGGCTCTACTAGGCTTGTGGCCGGGCTCCG -GCCGCGCTCC 150
61 GGGGCCCTTTTCTAGAGGGGCTCTACTAGGCTTGTGGCCGGGCTCCGCGCCCTCC 120
151 TCTGGATTGCCCCGAACACCGTGGTACTGTTGTTGCCGACAGAGAGCGCTGGTGGTG 210
|||||
121 TCTGGATTGCCCCGAACACCGTGGTACTGTTGTTGCCGACAGAGAGCGCTGGTGGTG 180
211 GAGCGAATGGGCGCATTCACCGGATCCTGGAGCCTGGTTGAACATCCTATCCCTGTG 270
181 GAGCGAATGGGCGCATTCACCGGATCCTGGAGCCTGGTTGAACATCCTATCCCTGTG 240
271 TTAGACCGGATCCGATATGTGAGAGTCTCAAGAAATTTGTCATCAACGTGCTGAGCAG 330
241 TTAGACCGGATCCGATATGTGAGAGTCTCAAGAAATTTGTCATCAACGTGCTGAGCAG 300
331 TCGGGCTGTGACTCTCGACAAATGTAATCTGCAAAATGATGAGTGCCTTTTACCTGGCATC 390
301 TCGGGCTGTGACTCTCGACAAATGTAATCTGCAAAATGATGAGTGCCTTTTACCTGGCATC 360
391 ATGAGACCTTACAAGCAAGCTAGCGTGTGAGAGACCTGAGTATGCCGTCAACCACTA 450
361 ATGAGACCTTACAAGCAAGCTAGCGTGTGAGAGACCTGAGTATGCCGTCAACCACTA 420
451 GGTCAACACCATGAGATCAGAGTCGGGCAACCTCTGTTGGACAAAGTGTCCGGGAA 510
421 GGTCAACACCATGAGATCAGAGTCGGGCAACCTCTGTTGGACAAAGTGTCCGGGAA 480
511 CGGAGATCCCTGAATGCCAGCATTTGTGATGCCATCAACCAAGCTGTGACTGTGGGGT 570
481 CGGAGATCCCTGAATGCCAGCATTTGTGATGCCATCAACCAAGCTGTGACTGTGGGGT 540
571 ATCCGCTGCCCTTCATTATGATCAAGATATCATGTGCCACCCCGGGTGAAGATCT 630
|||||

Db 541 ATCCGCTGCCCTCCGTTATGATGATCAAGATATCATGTGCCACCCGGGTGAAGAGCTCT 600
QY 631 ATGCAGATCGAGGTGGAGGACGAGCGGGAACGGGCGACAGTTCTAGAGTCTGAGGGG 690
Db 601 ATCCAGATCGAGGTGGAGGACGAGCGGGAACGGGCGACAGTTCTAGAGTCTGAGGGG 660
QY 691 ACCCGAGAGTCGGCCATCAATGTGGCAGAAGGAGAAGAACAGGCCCATCTGGGCTCC 750
Db 661 A-CCGAGAGTCGGCCATCAATGTGGCAGAAGGAGAAGAACAGGCCCATCTGGGCTCC 719
QY 751 GAAGCAAAAGGCTGAACAGATTAATTCAGGACGACGAGAAAGGCCCATGCTGCTTCT- GGC 809
Db 720 GAAGCAAAAGGCTGAACAGATTAATTCAGGACGACGAGAAAGGCCCATGCTGCTTCTG 779
QY 810 GAAGCAAAAGGCTGAACAGATTAATTCAGGACGACGAGAAAGGCCCATGCTGCTTCTG 868
Db 780 GAAGCAAAAGGCTGAACAGATTAATTCAGGACGACGAGAAAGGCCCATGCTGCTTCTG 839
QY 869 ATGGAGATGACGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 903
Db 840 ATGGAGATGACGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 873

RESULT 15
BI862105
LOCUS 920 bp mRNA linear EST 10-OCT-2001
DEFINITION 60339095F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5400023 5',
ACCESSION BI862105
VERSION BI862105.1 GI:16002852
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 920)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cgraphs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jnl.gov
Place: LHAM12019 row: n column: 24
High quality sequence stop: 918.
Location/Qualifiers
1. 920
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5400023"
/clone_1ib="NIH_MGC_87"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="PH10B (phage-resistant)"
/note="Organ: breast; Vector: pCMV-SPORT5; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH-MGC Library."
BASE COUNT 241 a 241 c 259 g 179 t
ORIGIN

Query Match 66.6%; Score 791.2; DB 10; Length 920;
Best Local Similarity 99.0%; Pred. No. 4e-183;
Matches 804; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 373 GTTCCTTACCTGCGCATGACCTTACAAAGCAAGCTACGCTGTGAGAGACCTGAG 432
|||||

Db 1 GTTCCTTACCTGCGCATGACCTTACAAAGCAAGCTACGCTGTGAGAGACCTGAG 60
QY 433 TATCCCTCCACCCGATGCTCAACAAACATGATTCAGAGCTCGGCAAACTCTCTNTG 492
Db 61 TATCCCTCCACCCGATGCTCAACAAACATGATTCAGAGCTCGGCAAACTCTCTCTG 120
QY 493 GACAAATCTTCCGGGAACGGAGATCCCTGAATCCAGATCTTGGATGCCATCAACAA 552
Db 121 GACAAATCTTCCGGGAACGGAGATCCCTGAATCCAGATCTTGGATGCCATCAACAA 180
QY 553 GCTGCTACTGCTGGGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 612
Db 181 GCTGCTACTGCTGGGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 613 CCCCAGTGAAGAGCTATGCAATGATGAGTGAGGACGAGCGGCGGAAGCGGCGACA 672
Db 241 CCCCAGTGAAGAGCTATGCAATGATGAGTGAGGACGAGCGGCGGAAGCGGCGACA 300
QY 673 GTTCTAGAGTCTGAGGGGACCCGAGAGTCCGCTCAATGATGAGTGAGGACGAGCGGCGACA 732
Db 301 GTTCTAGAGTCTGAGGGGACCCGAGAGTCCGCTCAATGATGAGTGAGGACGAGCGGCGACA 360
QY 733 GCCAGATCTCTGGCTCTCGAAGCAGAAAGGCTGAACAGATTAATTCAGCAGCAGAGAG 792
Db 361 GCCAGATCTCTGGCTCTCGAAGCAGAAAGGCTGAACAGATTAATTCAGCAGCAGAGAG 420
QY 793 GCCAGTGCAGTCTGGGGAAGGCGCAAGGCTTAAGCTGAGCTGATTCGATCTGCTGCTGCA 852
Db 421 GCCAGTGCAGTCTGGGGAAGGCGCAAGGCTTAAGCTGAGCTGATTCGATCTGCTGCTGCA 480
QY 853 GCTCTGACACACATTAATGAGATGACAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 911
Db 481 GCTCTGACACACATTAATGAGATGACAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 912 CAGCGCTTCTCCAAACTGGCCAAAGACTCCACATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTG 971
Db 541 CAGCGCTTCTCCAAACTGGCCAAAGACTCCACATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 972 CGATGTCACACAGATGCTGGGCTGAGGCGCATGAGGCTATGAGAGCCCTCACCAAGGCCCC 1031
Db 601 CGATGTCACACAGATGCTGGGCTGAGGCGCATGAGGCTATGAGAGCCCTCACCAAGGCCCC 660
QY 1032 AGTGCCAGGAGCTCCAGACTCTCCAGTGGGAGGACGAGAGATGCTCCAGGCTACAGA 1091
Db 661 AGTGCCAGGAGCTCCAGACTCTCCAGTGGGAGGACGAGAGATGCTCCAGGCTACAGA 720
QY 1092 TGCAGATTTGATGAGGAGACTTGTGAGATCAAGATGAGTGTAGTGGAGCTGGCTTGGCC 1151
Db 721 TGCAGATTTGATGAGGAGACTTGTGAGATCAAGATGAGTGTAGTGGAGCTGGCTTGGCC 780
QY 1152 AGGAGTCTGGGGGACAGGAAGCAGATTTC 1183
Db 781 AGGAGTCTGGGGGACAGGAAGCAGATTTC 812

Search completed: September 22, 2002, 14:49:21
Job time: 5287 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 18:32:18 : Search time 14.69 Seconds

(without alignments)
938.336 Million cell updates/sec

Title: US-09-898-216-1

Perfect score: 356

Sequence: 1 MLARAARGHMGCPFAEGLSTG.....RDVQGTDAKXDELDKRYKMS 356

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	2.5	160	VP12_BPPRD	P17637 bacterioph
2	8	2.2	343	UL14_HCMVA	P16756 human cytom
3	7	2.0	92	ST05_RAT	P50231 rattus norv
4	7	2.0	134	ATPE_RHORI	P05442 rhodospirill
5	7	2.0	158	RL29_HUMAN	P47914 homo sapien
6	7	2.0	164	RPOE_BACHD	O9666 bacillus ha
7	7	2.0	166	KAC5_BOVIN	P04262 bos taurus
8	7	2.0	180	YFA2_ECOLI	P76471 escherichia
9	7	2.0	194	UREF_HAEIN	P44395 haemophilus
10	7	2.0	228	LOLD_VIBCH	P57066 vibrio chol
11	7	2.0	232	RADC_BACHD	O9884 bacillus ha
12	7	2.0	279	FIXA_RHILP	O0559 rhizobium l
13	7	2.0	281	REPL_BPP1	P19654 bacterioph
14	7	2.0	307	OTC_ARCFU	O29013 archaeoglob
15	7	2.0	344	IN37_SPIOL	P23525 spinacia ol
16	7	2.0	347	RECA_CHRYI	O32377 chromatium
17	7	2.0	372	TOLA_HAEIN	P44628 haemophilus
18	7	2.0	381	DUS6_HUMAN	O16828 homo sapien
19	7	2.0	381	DUS6_MOUSE	O96db1 mus musculi
20	7	2.0	381	DUS6_RAT	O64446 rattus norv
21	7	2.0	381	YE88_MYCTU	P71768 mycobacteri
22	7	2.0	421	TOLA_ECOLI	P19934 escherichia
23	7	2.0	422	KICW_HUMAN	O9c075 homo sapien
24	7	2.0	426	FLT1_DROME	O61491 drosophila
25	7	2.0	432	TTC_ECOLI	P22357 escherichia
26	7	2.0	433	Z001_YEAST	P32527 saccharomyc
27	7	2.0	435	MDHP_SPIOL	P52426 spinacia ol
28	7	2.0	448	TBA4_HUMAN	O55215 homo sapien
29	7	2.0	475	Z131_HUMAN	P52739 homo sapien
30	7	2.0	478	PKKH_ARATH	P50318 arabidopsis
31	7	2.0	480	PKKH_WHEAT	P12782 triticum ae
32	7	2.0	517	UZAF_SCHPO	P36629 schizosacch
33	7	2.0	517	YC09_CAMJE	O9pn86 campylobact

34	7	2.0	553	1	FXC1_HUMAN	O12948 homo sapien
35	7	2.0	553	1	FXC1_MOUSE	O61572 mus musculi
36	7	2.0	570	1	IEP7_CAEBL	P30642 caenorhabdit
37	7	2.0	580	1	RGP1_XENLA	O13066 xenopus lae
38	7	2.0	587	1	RGP1_HUMAN	P46061 homo sapien
39	7	2.0	589	1	RGP1_MOUSE	P46061 mus musculi
40	7	2.0	655	1	SCAG_MOUSE	O9wu39 mus musculi
41	7	2.0	693	1	SPAL_MOUSE	P46062 mus musculi
42	7	2.0	704	1	HS90_SCHRO	P41887 schizosacch
43	7	2.0	706	1	SEM2_MACMO	O95196 macaca mula
44	7	2.0	739	1	RELA_STRDO	O54089 streptococc
45	7	2.0	891	1	MAZ3_SCHCO	P37937 schizophy11

ALIGNMENTS

```

RESULT 1
ID      VP12_BPPRD      STANDARD:      PRT:      160 AA.
AC      P17637;
DT      01-AUG-1990 (rel. 15, Created)
DT      01-AUG-1990 (rel. 15, Last sequence update)
DT      01-DEC-1992 (rel. 24, Last annotation update)
DE      Single-stranded DNA-binding protein P12.
GN      XII.
OS      Bacteriophage PRD1.
OC      Viruses; dsDNA viruses, no RNA stage; Tectiviridae; Tectivirus.
OX      NCBI_TaxID=10658;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=90152379; PubMed=2695404;
RA      Pakula T.M., Savilabti H., Bamford D.H.;
RT      "The organization of the right-end early region of bacteriophage PRD1
RT      genome."
RL      Gene 85:53-58(1989).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=90202706; PubMed=2180910;
RA      Gerendasy D., Ito J.;
RT      "Nucleotide sequence and transcription of the right early region of
RT      bacteriophage PRD1."
RL      J. Bacteriol. 172:1889-1898(1990).
CC      -I- FUNCTION: P12 HAS A REGULATORY EFFECT ON PHAGE DNA METABOLISM
CC      AND TRANSCRIPTION OF EARLY GENES. BINDS TO SINGLE-STRANDED DNA.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL: M69077; AAA32476.1; -
DR      EMBL: M33428; AAA32447.1; -
DR      EMBL: M30146; AAA32453.1; -
DR      PIR: J00186; J00186.
DR      PIR: A35148; A35148.
KW      Early protein; DNA-binding.
SQ      SEQUENCE 160 AA: 16650 MW: 461004D25FC9F470 CRC64:

```

Query Match 2.5%; Score 9; DB 1; Length 160;
Best local Similarity 100.0%; Pred. No. 0.23;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 249 AKAKAKAEA 257
|||||
DB 152 AKAKAKAEA 160

RESULT 2

```

UL14_HCMVA          STANDARD:          PRT:          343 AA.
AC P16756;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Hypothetical protein UL14.
GN UL14.
OS Human cytomegalovirus (strain AD169).
OC Herpesviridae; Cytomegalovirus.
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10360;
RN [1]
RP MEDLINE=90269039; PubMed=2161319;
RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
RA Hornell T., Hutchinson C.A. III, Kouzarides T., Martignetti J.A.,
RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RT "Analysis of the protein-coding content of the sequence of human
RT cytomegalovirus strain AD169."
RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X17403; CAA35447.1; -.
DR PIR: S09777; S09777.
DR HSSP: P04002; IATF.
KW Hypothetical protein.
FT CARBOHYD 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 343 AA: 38566 MW: F4A181A14FDAL98 CRC64;

Query Match          2.2%; Score 8; DB 1; Length 343;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 240 AAGEASAV 247
Db 323 AAGEASAV 330

RESULT 3
SY05_RAT          STANDARD:          PRT:          92 AA.
AC P50231;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Small inducible cytokine A5 precursor (T-cell specific RANTES protein)
DE (SIS-delta).
GN SCYA5.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=LONG EVANS; TISSUE=Lung;
RA Jones M.L., Shanley T.P., Ward P.A.;
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CHEMOKINE RECEPTOR FOR BLOOD MONOCYTES, MEMORY T HELPER
CC CELLS AND EOSINOPHILS. CAUSES THE RELEASE OF HISTAMINE FROM
CC BASOPHILS AND ACTIVATES EOSINOPHILS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE INTERCINE BETA FAMILY (SMALL CYTOKINE
CC C-C) (CHEMOKINE CC).

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U06436; AAA96499.1; -.
DR HSSP: P13501; IRTN.
DR InterPro: IPR001811; Chemokine_IL8.
DR InterPro: IPR000827; Small_cytokine_CC.
DR Pfam: PF00048; IL8; 1.
DR SMART: SM00199; SCY; 1.
DR PROSITE: PS00472; SMALL_CYTOKINES_CC; 1.
KW Cytokine; Chemotaxis; T-cell; Signal; Inflammatory response.
FT SIGNAL 1 24
FT CHAIN 25 92
FT DISULFD 34 58 SMALL INDUCIBLE CYTOKINE A5.
FT DISULFD 35 74 BY SIMILARITY.
SQ SEQUENCE 92 AA: 10170 MW: B4F8EC2B4208ABC6 CRC64;

OY 272 AAASLTIV 278
Db 5 AAASLTIV 11

Query Match          2.0%; Score 7; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 272 AAASLTIV 278
Db 5 AAASLTIV 11

RESULT 4
ATPE_RHORU          STANDARD:          PRT:          134 AA.
AC P05442;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE ATP synthase epsilon chain (EC 3.6.3.14) (ATP synthase F1 sector
DE epsilon subunit).
GN ATPC.
OS Rhodospirillum rubrum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillaceae;
OC Rhodospirillum.
OX NCBI_TaxID=1085;
RN [1]
RP SEQUENCE FROM N.A.
RP MEDLINE=85251588; PubMed=2861810;
RA Falk G., Hampe A., Walker J.E.;
RT "Nucleotide sequence of the Rhodospirillum rubrum atp operon.";
RL Biochem. J. 228:391-407(1985).
CC -1- FUNCTION: Produces ATP from ADP in the presence of a proton
CC gradient across the membrane.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(in) = ADP + phosphate +
CC H(+)(out).
CC -1- SUBUNIT: F-type atpases have 2 components, CF(1) - the catalytic
CC core - and CF(0) - the membrane proton channel. CF(1) has five
CC subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0)
CC has three main subunits: A, B and C.
CC -1- SIMILARITY: BELONGS TO THE ATPASE EPSILON CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X02499; CAB97256.1; -.
DR PIR: S08584; PMOFE.

```

DR HSPF: P00832; IBSN.
 DR InterPro: IPR001469; ATP-synt_DE.
 DR Pfam: PF00401; ATP-synt_DE.1.
 DR Pfam: PF02823; ATP-synt_DE.N: 1.
 DR ProDom: PD000944; ATP-synt_DE.1.
 DR Hydrolase: ATP synthetase; Cf(1); Hydrogen ion transport.
 KW SEQUENCE 134 AA; 14307 MW; 8905084DF31C9222 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 134;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 229 SEAEKAE 235
 DB 111 SEAEKAE 117

RESULT 5
 ID RL29_HUMAN STANDARD; PRT; 158 AA.
 AC P47914:
 DT 01-FEB-1996 (rel. 33, Created)
 DT 01-FEB-1996 (rel. 33, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE 60S ribosomal protein L29 (Cell surface heparin binding protein HRP).
 GN RPL29.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Law P.T.W., Tsui S.K.W., Lee C., Waye M.M.Y.;
 RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=96180309; PubMed=8597591;
 RA Law P.T., Tsui S.K., Lam W.Y., Luk S.C., Hwang D.M., Liew C.C.,
 RA Lee C.Y., Fung K.P., Waye M.M.;
 RL "A novel cDNA encoding a human homologue of ribosomal protein L29";
 RL Biochim. Biophys. Acta 1305:105-108(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=96216099; PubMed=8662616;
 RA Liu S., Smith S.E., Julian J., Rohde L.H., Karin N.J., Carson D.D.;
 RT "cDNA cloning and expression of HRP, a novel cell surface heparan
 sulfate/heparin-binding protein of human uterine epithelial cells and
 cell lines.";
 RL J. Biol. Chem. 271:11817-11823(1996).
 CC -1- SIMILARITY: BELONGS TO THE L29E FAMILY OF RIBOSOMAL PROTEINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U10248; AAC50459.1; -
 DR EMBL: 249148; CAAB9008.1; -
 DR EMBL: U49083; AAC50647.1; -
 DR MIM: 601832; -
 DR InterPro: IPR002673; Ribosomal_L29e.
 DR Pfam: PF01779; Ribosomal_L29e.1.
 DR ProDom: PD010314; Ribosomal_L29e.1.
 KW Ribosomal protein; Repeat; Heparin-binding.
 FT INIT MET 0 R -> A (TN REF. 2)
 FT CONFLICT 119 119 BY SIMILARITY.
 SQ SEQUENCE 158 AA; 17621 MW; 8FCC38682E5DDF28 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 158;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 249 AKAKAKA 255
 DB 122 AKAKAKA 128

RESULT 6
 ID RPOE_BACHD STANDARD; PRT; 164 AA.
 AC O9K6D6:
 DT 16-OCT-2001 (rel. 40, Created)
 DT 16-OCT-2001 (rel. 40, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE Probable DNA-directed RNA polymerase delta subunit (RNAP delta
 factor).
 GN RPOE OR BH3793.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 CC -1- FUNCTION: PARTICIPATES IN BOTH THE INITIATION AND RECYCLING PHASES
 CC OF TRANSCRIPTION. IN THE PRESENCE OF THE DELTA SUBUNIT, RNAP
 CC DISPLAYS AN INCREASED SPECIFICITY OF TRANSCRIPTION, A DECREASED
 CC AFFINITY FOR NUCLEIC ACIDS, AND AN INCREASED EFFICIENCY OF RNA
 CC SYNTHESIS BECAUSE OF ENHANCED RECYCLING (BY SIMILARITY).
 CC -1- SUBUNIT: RNAP IS COMPOSED OF A CORE OF 2 ALPHA, A BETA AND A
 CC BETA' SUBUNITS. THE CORE IS ASSOCIATED WITH A DELTA SUBUNIT AND
 CC ONE OF SEVERAL SIGMA FACTORS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE RPOE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AP001520; BAB07512.1; -
 DR DNA-directed RNA polymerase; Transcription; Complete proteome.
 KW DOMAIN 104 ASP/GLU-RICH (ACIDIC).
 FT SEQUENCE 164 AA; 19213 MW; 26E91F66DE67CA CRC64;

Query Match 2.0%; Score 7; DB 1; Length 164;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 250 KAKAKAE 256
 DB 98 KAKAKAE 104

RESULT 7
 ID K2C5_BOVIN STANDARD; PRT; 166 AA.
 AC P04262:
 DT 20-MAR-1987 (rel. 04, Created)
 DT 20-MAR-1987 (rel. 04, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)

```

DE Keratin, type II cytoskeletal 68 kDa, component IB (Fragment).
CC Bos taurus (Bovine).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
CC Bovidae; Bovinae; Bos.
OX NCBI_TaxId=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85128114; PubMed=6084625;
RA Jorcano J.L., Franz J.K., Franke W.M.;
RT "Amino acid sequence diversity between bovine epidermal cytokeratin
RT polypeptides of the basic (type II) subfamily as determined from cDNA
RT clones."
RL Differentiation 28:155-163(1984).
CC -1- SUBUNIT: HETEROTETRAMER OF TWO TYPE I AND TWO TYPE II KERATINS.
CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF CYTOSKELETAL AND
CC MICROFILAMENTAR KERATIN, I (ACIDIC) AND II (NEUTRAL TO BASIC)
CC (40-55 AND 56-70 KILODALTONS, RESPECTIVELY).
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: K03534; AAA30601.1; -.
CC PIR: A02948; KR802B.
DR HSP: P10968; JMG8.
DR InterPro: IPR001664; IF.
DR Pfam: PF00038; Filament; 1.
DR PROSITE: PS00226; IF: 1.
KW Intermediate filament; Coiled coil; Keratin.
FT NON_TER 1 1
FT DOMAIN <1 41 ROD.
FT DOMAIN 42 166 TAIL.
FT DOMAIN <1 41 COIL_2B.
OX SEQUENCE 166 AA; 15505 MW; 4BFE495A7C6A8E0 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 166;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 331 LSSGSSR 337
DB 145 LSSGSSR 151

RESULT 8
YFAZ_ECOLI
ID YFAZ_ECOLI STANDARD: PRT: 180 AA.
AC P76471:
DT 15-DEC-1998 (rel. 37, Created)
DT 15-DEC-1998 (rel. 37, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE Hypothetical protein yfaz precursor.
GN YFAZ OR B2250.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxId=562;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."

```

```

RL Science 277:1453-1474(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000315; AAC75310.1; ALT_INIT.
DR Ecogene; B614088; yfaz.
KW Hypothetical protein; Signal; Complete proteome.
FT SIGNAL 1 21
FT CHAIN 22 180 POTENTIAL.
FT SIGNAL 1 21
OX SEQUENCE 180 AA; 18610 MW; E0836316B9AD5797 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 180;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 PDSLSSG 334
DB 121 PDSLSSG 127

RESULT 9
UREF_HAEIN
ID UREF_HAEIN STANDARD: PRT: 194 AA.
AC P44395:
DT 01-NOV-1995 (rel. 32, Created)
DT 01-NOV-1995 (rel. 32, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE Urease accessory protein uref.
GN UREF OR H10537.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxId=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=9350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd."
RL Science 269:496-512(1995).
CC -1- FUNCTION: PROBABLY FACILITATES NICKEL INCORPORATION.
CC -1- SIMILARITY: BELONGS TO THE UREF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U32736; AAC22195.1; -.
DR TIGR; H10537; -.
DR InterPro: IPR002639; UREF.
DR Pfam: PF01730; Uref: 1.
DR ProDom: PD004961; Uref: 1.
KW Nickel; Complete proteome.

```

SQ SEQUENCE 194 AA; 21960 MW; B38F134AE3DCDD0B5 CRC64;

Query Match
Best Local Similarity 2.0%; Score 7; DB 1; Length 194;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 308 MVAQAMG 314
|||||||
Db 105 MVAQAMG 111

RESULT 10
LOAD_VIBCH STANDARD; PRT: 228 AA.

AC P57066;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lipoprotein releasing system ATP-binding protein 10LD.
GN 10LD OR VC1883.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Swinn M.L.,
RA Dodson R.J., Hatt D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolova M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Uterback T., Fleischmann R.D., Nieman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae.";
RL Nature 406:477-483(2000).
CC -1- FUNCTION: PART OF AN ATP-DEPENDENT TRANSPORT SYSTEM RESPONSIBLE
FOR THE RELEASE OF LIPOPROTEINS TARGETED TO THE OUTER MEMBRANE
FROM THE INNER MEMBRANE. SUCH A RELEASE IS DEPENDENT OF THE
SORTING-SIGNAL (ABSENCE OF AN ASP AT POSITION 2 OF THE MATURE
LIPOPROTEIN) AND OF LOLA (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. 10LD SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE004263; AAF95031.1; -;
DR TIGR: VC1883; -;
DR InterPro: IPR003593; AAA.
DR InterPro: IPR003439; ABC_transportr.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00005; ABC_tran; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; Transport; Complete proteome.
FT NP_BIND 41 48
FT BIND 41 48
FT SEQUENCE 228 AA; 24690 MW; 315E52CB775E61E6 CRC64;

Query Match
Best Local Similarity 2.0%; Score 7; DB 1; Length 228;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 249 AKAKAKA 255
|||||||

Db 119 AKAKAKA 125

RESULT 11
RADCC_BACBD STANDARD; PRT: 232 AA.

AC Q9K8H4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA repair protein radC homolog.
GN RADCC OR BH3032.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=66665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- FUNCTION: INVOLVED IN DNA REPAIR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE RADCC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AP001517; BAB06751.1; -;
DR InterPro: IPR000445; HHH.
DR InterPro: IPR001405; RadC.
DR Pfam: PF00633; HHH; 1.
DR ProDom: PD007415; RadC; 1.
DR PROSITE: PS01302; RADCC; 1.
KW DNA repair; Complete proteome.
FT SEQUENCE 232 AA; 26403 MW; 4542153B5C6010D2 CRC64;

Query Match
Best Local Similarity 2.0%; Score 7; DB 1; Length 232;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 152 SLMASIV 158
|||||||
Db 151 SLMASIV 157

RESULT 12
FIXA_RHILP STANDARD; PRT: 279 AA.

AC Q05559;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Fixa protein.
GN FIXA.
OS Rhizobium leguminosarum (biovar phaseolli).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=385;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93379046; PubMed=8369342;
RA Michiels J., Vanderleyden J.;

```

RT      "Cloning and sequence of the Rhizobium leguminosarum biovar phaseoli
RL      fixa gene."
RL      Blochm. Biophys. Acta 1144:232-233(1993).
CC      -1- FUNCTION: MAY PLAY A ROLE IN A REDOX PROCESS INVOLVED IN NITROGEN
CC      FIXATION.
CC      -1- SUBUNIT: FIXA AND FIXB FORM A HETERODIMER (PROBABLE).
CC      -1- SIMILARITY: BELONGS TO THE ETF BETA-SUBUNIT / FIXA FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: L11081; AAA02979.1; -.
DR      HSSP: P38117; IEFV.
DR      InterPro: IPR000049; ETF_beta.
DR      Pfam: PF01012; ETF_beta.1.
DR      ProDom: PD003528; ETF_beta.1.
DR      PROSITE: PS01065; ETF_BETA.1.
KW      Electron transport; Nitrogen fixation.
SQ      SEQUENCE 279 AA; 30083 MW; 6ACB3EABF20D1CE8 CRC64;

Query Match          2.0%; Score 7; DB 1; Length 279;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      231 AKKAOI 237
        |||||
Db      240 AKKAOI 246

RESULT 13
REPL_Bp1      STANDARD:      PRT:      281 AA.
AC      P19654;
DT      01-FEB-1991 (Rel. 17, Created)
DT      01-FEB-1991 (Rel. 17, Last sequence update)
DT      01-FEB-1991 (Rel. 17, Last annotation update)
DE      Repl. protein.
CN      REPL.
OS      Bacteriophage P1.
OC      Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX      NCBI_TaxID=10678;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE-69293847; Pubmed-2661831;
RA      Hansen E.B.;
RT      "Structure and regulation of the lytic replicon of phage P1.";
RL      J. Mol. Biol. 207:135-149(1989).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE-69293846; Pubmed-2661830;
RA      Sternberg N.; Cohen G.;
RT      "Genetic analysis of the lytic replicon of bacteriophage P1. II.
RT      Organization of replicon elements.";
RL      J. Mol. Biol. 207:111-133(1989).
CC      -1- FUNCTION: ESSENTIAL FOR L-REPLICON FUNCTION. NECESSARY FOR
CC      INITIATION OF DNA REPLICATION FROM THE L-REPLICON (PROBABLE).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: X15639; CAA33662.1; -.
DR      EMBL: X15638; CAA33656.1; -.

```

```

DR      EMBL: X15638; CAA33657.1; ALT_INIT.
DR      PIR: S04263; RLBP1.
KW      DNA replication.
SQ      SEQUENCE 281 AA; 30857 MW; F3C1D033BD756138 CRC64;

Query Match          2.0%; Score 7; DB 1; Length 281;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      249 AKAKAK 255
        |||||
Db      198 AKAKAK 204

RESULT 14
OTC_ARCFU      STANDARD:      PRT:      307 AA.
ID      OTC_ARCFU
AC      O29013;
DT      15-JUL-1998 (Rel. 36, Created)
DT      15-JUL-1998 (Rel. 36, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Ornithine carbamoyltransferase (EC 2.1.3.3) (OTCase).
GN      ARGF OR ARL255.
OS      Archaeoglobus fulgidus.
OC      Archaeae; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC      Archaeoglobus.
OX      NCBI_TaxID=2234;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX      MEDLINE-98049343; Pubmed-9389475;
RA      Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA      Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA      Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA      Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA      Kinkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA      Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA      Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA      Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA      Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA      Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA      Venter J.C.;
RT      "The complete genome sequence of the hyperthermophilic, sulphate-
RT      reducing archaeon Archaeoglobus fulgidus.";
RL      Nature 390:364-370(1997).
CC      -1- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine -> phosphate
CC      + L-citrulline.
CC      -1- PATHWAY: ARGinine BIOSYNTHESIS (POTENTIAL).
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC      -1- SIMILARITY: BELONGS TO THE ARCUSES/OTCASES FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: AE001017; AAB89987.1; -.
DR      HSSP: Q51742; IAI5.
DR      TIGR: AF1255; -.
DR      InterPro: IPR002029; Carbmyltransf_asor.
DR      Pfam: PF00185; OTCase.1.
DR      PRINTS: PR00100; AOTCASE.
DR      PROSITE: PS00097; CARBAMOYLTRANSFERASE; FALSE_NEG.
KW      Transferase; Arginine biosynthesis; Complete proteome.
SQ      SEQUENCE 307 AA; 34913 MW; DB043BC5A40EB5D CRC64;

Query Match          2.0%; Score 7; DB 1; Length 307;

```

Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 ILAALT 266
|11111|
DB 166 ILAALT 172

RESULT 15

IN37_SPIOL STANDARD; PRT; 344 AA.
ID IN37_SPIOL
AC P23525;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE 37 kDa inner envelope membrane protein, chloroplast precursor (E37).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91146572; PubMed=1997321;
RA Drees-Werlingloer U., Fischer K., Wachter E., Link T.A.,
RA Fluegge U.J.;
RT "CDNA sequence and deduced amino acid sequence of the precursor of
RT the 37-kDa inner envelope membrane polypeptide from spinach
RT chloroplasts. Its transit peptide contains an amphiphilic alpha-helix
RT as the only detectable structural element.";
RT Eur. J. Biochem. 195:361-368(1991).
RN [2]
RP SEQUENCE OF 120-141 AND 236-240.
RX MEDLINE=91348205; PubMed=1879527;
RA Block M.A., Joyard J., Douce R.;
RT "Purification and characterization of E37, a major chloroplast
RT envelope protein.";
RL FEBS Lett. 287:167-170(1991).
CC -1- FUNCTION: NOT YET KNOWN. ONE OF THE MAJOR PROTEINS OF THE INNER
CC MEMBRANE OF THE CHLOROPLAST ENVELOPE.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST INNER MEMBRANE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X56963; CAA40283.1; -.
DR PIR; S14409; S14409.
DR InterPro: IPR000051; SAM_bind.
DR InterPro: IPR004033; ubiE_COO5_methyltransf.
DR Pfam: PF01209; ubiE_methyltran; 1.
KW Chloroplast; transit peptide; inner membrane.
FT TRANSIT 1 344 CHLOROPLAST.
FT CHAIN 22 344 37 KDA INNER ENVELOPE MEMBRANE PROTEIN.
SQ SEQUENCE 344 AA: 38976 MW: 555E53242B297D7C CRC64;

Query Match 2.0%; Score 7; DB 1; Length 344;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 LAKAKAK 254
|111111|
DB 154 LAKAKAK 160

0 1 2 3 4 5 6 7 8 9

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 18:30:07 ; Search time 20.37 Seconds
(without alignments)
1679.323 Million cell updates/sec

Title: US-09-898-216-1

Perfect score: 356

Sequence: 1 MLRAARHGMPFAEGLSTG.....RDVQGTDAKDDELDRVAMS 356

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	152	42.7	357	2 T02246	hypothetical prote
2	14	3.9	334	2 T21562	hypothetical prote
3	13	3.7	515	2 T05863	hypothetical prote
4	11	3.1	311	2 B71689	hypothetical prote
5	11	3.1	312	2 D97756	hypothetical prote
6	11	3.1	354	2 T39599	conserved hypothet
7	9	2.5	160	2 A35148	early protein p12
8	9	2.5	374	2 T18493	hypothetical prote
9	8	2.2	248	2 E87425	transcription regu
10	8	2.2	263	2 G83163	probable amino aci
11	8	2.2	343	2 S09777	hypothetical prote
12	8	2.2	398	2 C96680	hypothetical prote
13	8	2.2	400	2 AG3259	hypothetical prote
14	8	2.2	477	2 T18801	hypothetical prote
15	8	2.2	493	2 D98209	hypothetical prote
16	8	2.2	493	2 AE3077	aldelyde dehydrat
17	8	2.2	658	2 T20042	aldelyde dehydrat
18	8	2.2	658	2 T28658	hypothetical prote
19	7	2.0	2124	2 T47015	polyketide synthas
20	7	2.0	47	2 I47015	antigen WC1 (stiml
21	7	2.0	95	2 D83325	hypothetical prote
22	7	2.0	95	2 B45519	variant surface g1
23	7	2.0	112	2 F71805	hypothetical prote
24	7	2.0	112	2 D64713	conserved hypothet
25	7	2.0	120	2 A45519	variant surface g1
26	7	2.0	121	2 E72617	hypothetical prote
27	7	2.0	134	1 PMORE	H+-transporting tw
28	7	2.0	135	2 C72539	hypothetical prote
29	7	2.0	136	2 F95868	conserved hypothet
29	7	2.0	159	2 S65784	ribosomal protein

30	7	2.0	164	2 A84124	DNA-directed RNA p
31	7	2.0	166	1 KRB02B	keratin, 68k type
32	7	2.0	175	2 S43815	hypothetical prote
33	7	2.0	183	2 C97295	stage V sporulatio
34	7	2.0	187	2 H64395	hypothetical prote
35	7	2.0	187	2 B91021	hypothetical prote
36	7	2.0	194	2 A85865	hypothetical prote
37	7	2.0	194	2 F64075	urase accessory p
38	7	2.0	207	2 F83148	hypothetical prote
39	7	2.0	209	2 T17294	hypothetical prote
40	7	2.0	210	2 JC4683	T-cell receptor ga
41	7	2.0	210	2 B75372	hypothetical prote
42	7	2.0	214	2 C82556	hypothetical prote
43	7	2.0	215	2 AE0793	conserved hypothet
44	7	2.0	228	2 A82147	ABC transporter, A
45	7	2.0	232	2 H84028	DNA repair protein

ALIGNMENTS

RESULT 1
T02246
hypothetical protein P1.11659_4 - human
C:Species: Homo sapiens (man)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 11-Jan-2000
C:Accession: T02246
R:lamerdin, J.E.; McCreedy, P.M.; Skowronski, E.; Adamson, A.W.; Burkhardt-Schultz, K.; Christensen, M.; Georgescu, A.; Avila, J.; Liu, S.; Altix, C.; Andreise, F.; Trank
Submitted to the EMBL Data Library, March 1998
A:Authors: Duarte, S.; Lucas, S.; Bruce, R.; Thomas, P.; Quan, G.; Kronmiller, B.; Ar
A:Description: Sequence analysis of a human P1 clone containing the XRC9 DNA repair
A:Reference number: Z14637
A:Accession: T02246
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-357 <LAMB>
A:Cross-references: EMBL:AC004472; NID:g2984582; PIDN:AC07983.1; PID:g2984585
C:Genetics:
A:Map position: 9
A:Insertions: 2/3; 48/3; 82/1; 101/3; 135/3; 194/3; 243/1; 269/3; 312/3
C:Superfamily: erythrocyte band 7 integral membrane protein

Query Match 42.7% Score 152; DB 2; Length 357;
Best Local Similarity 100.0% Pred. No. 4.9e-142;
Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 VEAERRRRATVLESEGTRESAINVAEGKKOQIILASEAEKAQIINQAAGEASAVLAKAKA 253
|||||
Db 195 VEAERRRRATVLESEGTRESAINVAEGKKOQIILASEAEKAQIINQAAGEASAVLAKAKA 254

QY 254 KAEAIRILAAALTOHNGDAASLTVAEOYVSASAKLAKDSNTILLPSNPGDVTSMAQAAM 313
|||||
Db 255 KAEAIRILAAALTOHNGDAASLTVAEOYVSASAKLAKDSNTILLPSNPGDVTSMAQAAM 314

QY 314 GYGALTKAPVPGPDSLSGSSSRDVGTDAS 345
|||||
Db 315 GYGALTKAPVPGPDSLSGSSSRDVGTDAS 346

RESULT 2
T21562
hypothetical protein F30A10.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T21562
R:Barlow, K.
Submitted to the EMBL Data Library, October 1996
A:Reference number: Z19442
A:Accession: T21562
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 1-334 <MIL>
 A:Cross-references: EMBL:Z81072; PIDN:CAB03018.1; GSPDB:GN00019; CESP:F30A10.5
 A:Experimental source: clone F30A10
 C:Genetics:
 A:Gene: CESP:F30A10.5
 A:Map position: 1
 A:introns: 16/2; 68/3; 124/3; 227/2
 C:Superfamily: erythrocyte band 7 integral membrane protein

Query Match 3.1%; Score 14; DB 2; Length 334;
 Best Local Similarity 100.0%; Pred. No. 1.5e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 40 FVPOEAMVERMG 53
 |||||||
 DB 47 FVPOEAMVERMG 60

RESULT 3
 T05863

hypothetical protein T29A15.70 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 13-Aug-1999
 C:Accession: T05863

R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voelt, M.; Robben, J.; Volckaert, G.; He
 submitted to the Protein Sequence Database, March 1999
 A:Reference number: 215455
 A:Accession: T05863

A:Molecule type: DNA
 A:Residues: 1-515 <BEV>
 A:Cross-references: EMBL:AL035602
 A:Experimental source: cultivar Columbia; BAC clone T29A15
 C:Genetics:

A:Map position: 4
 A:introns: 52/2; 133/3; 183/3; 203/1; 234/1; 267/1; 297/3; 317/3; 394/1; 451/3
 A:Note: T29A15.70

Query Match 3.7%; Score 13; DB 2; Length 515;
 Best Local Similarity 100.0%; Pred. No. 0.00021;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 128 QLAQTMRSGLK 140
 |||||||
 DB 153 QLAQTMRSGLK 165

RESULT 4
 B71689

hypothetical protein RP328 - Rickettsia prowazekii
 C:Species: Rickettsia prowazekii
 C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
 C:Accession: B71689

R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Slicheritz-Ponten, T.; Alsmark, U
 Nature 396, 133-140, 1998
 A:title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
 A:Reference number: A71630; MUID:99039499

A:Accession: B71689
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-311 <AND>

A:Cross-references: GB:AJ235271; GB:AJ235269; NID:g3868717; PIDN:CA14788.1; PID:g386088
 A:Experimental source: strain Madrid E
 C:Genetics:

A:Gene: RP328
 C:Superfamily: erythrocyte band 7 integral membrane protein

Query Match 3.1%; Score 11; DB 2; Length 311;
 Best Local Similarity 100.0%; Pred. No. 0.013;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 127 TOLAQTMRSE 137
 |||||||
 DB 109 TOLAQTMRSE 119

RESULT 5
 D97756

hypothetical protein RC0452 (imported) - Rickettsia conorii (strain Malish 7)

C:Species: Rickettsia conorii
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001
 C:Accession: D97756

R:ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fourmter, P.E.; Barbe, V.; Samson, D.;
 Science 293, 2093-2098, 2001

A:title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
 A:Reference number: A97700; MUID:21442074; PMID:11557893

A:Accession: D97756
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-312 <KUR>
 A:Cross-references: GB:AE006914; PIDN:ANL02990.1; PID:g15619524; GSPDB:GN00173
 C:Genetics:
 A:Gene: RC0452
 C:Superfamily: erythrocyte band 7 integral membrane protein

Query Match 3.1%; Score 11; DB 2; Length 312;
 Best Local Similarity 100.0%; Pred. No. 0.013;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 127 TOLAQTMRSE 137
 |||||||
 DB 109 TOLAQTMRSE 119

RESULT 6
 T39599

conserved hypothetical protein SPBC16G5.07c - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Feb-2000
 C:Accession: T39599

R:lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.
 submitted to the EMBL Data Library, May 1998

A:Reference number: Z21866
 A:Accession: T39599
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A:Residues: 1-354 <LYN>
 A:Cross-references: EMBL:AL023554; PIDN:CA119027.1; GSPDB:GN00067; SPDB:SPBC16G5.07c
 A:Experimental source: strain 97zh-; cosmid c16G5

C:Genetics:
 A:Gene: SPDB:SPBC16G5.07c
 A:Map position: 2
 A:introns: 72/2; 265/2
 C:Superfamily: erythrocyte band 7 integral membrane protein

Query Match 3.1%; Score 11; DB 2; Length 354;
 Best Local Similarity 100.0%; Pred. No. 0.014;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 111 DPKASYGVED 121
 |||||||
 DB 126 DPKASYGVED 136

RESULT 7
 A35148

early protein p12 - phage PRD1

C:Species: phage PRD1
 C:Date: 03-Aug-1990 #sequence_revision 03-Aug-1990 #text_change 08-Oct-1999
 C:Accession: A35148; J00186

R:gerendasy, D.; Ito, J.
 J. Bacteriol. 172, 1889-1898, 1990

A:Title: Nucleotide sequence and transcription of the right early region of bacteriophage
A:Reference number: A35148; MUID:90202706
A:Accession: A35148
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-160 <GER>
A:Cross-references: GB:M30146; NID:g215762; PIDN:AAA32453.1; PID:g215763
R:Palula, T.M.; Savillan, H.; Bamford, D.H.
Gene 85, 53-58, 1989
A:Title: The organization of the right-end early region of bacteriophage PRD1 genome.
A:Reference number: JQ0186; MUID:90152379
A:Accession: JQ0186
A:Molecule type: DNA
A:Residues: 1-160 <PAK>
A:Cross-references: GB:M3428; NID:g215747; PIDN:AAA32447.1; PID:g215748
C:Keywords: DNA binding

Query Match 2.5%; Score 9; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 249 AKAKAKAA 257
DB 152 AKAKAKAA 160

RESULT 8
T18493
hypothetical protein C0800w - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
R:Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z18935
A:Accession: T18493
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-374 <LAW>
A:Cross-references: EMBL:Z98551; NID:e1331903; PID:e1331914; PIDN:CAB11132.1
C:Genetics:
A:Map position: 3
A:Note: C0800w

Query Match 2.5%; Score 9; DB 2; Length 374;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 TMRSELGKL 141
DB 174 TMRSELGKL 182

RESULT 9
E87425
transcription regulator, GntR family [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 30-Sep-2001
C:Accession: E87425
R:Merz, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gelfand, M.L.; Haft, D.H.; Kolton
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: E87425
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-248 <STO>
A:Cross-references: GB:AE005673; NID:g13422783; PIDN:AAK23401.1; GSPDB:GNC0148
C:Genetics:

A:Gene: CC1420
C:Superfamily: pyruvate dehydrogenase complex repressor

Query Match 2.2%; Score 8; DB 2; Length 248;
Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 256 EAIRILAA 263
DB 65 EAIRILAA 72

RESULT 10
G83163
probable amino acid binding protein PA3865 [imported] - Pseudomonas aeruginosa (strain
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: G83163
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337
A:Accession: G83163
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-263 <STO>
A:Cross-references: GB:AE004803; GB:AE004091; NID:g9950035; PIDN:AA607252.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3865
C:Superfamily: lysine-arginine-ornithine-binding protein

Query Match 2.2%; Score 8; DB 2; Length 263;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 244 ASAVLAKA 251
DB 151 ASAVLAKA 158

RESULT 11
S09777
hypothetical protein ULI4 precursor - human cytomegalovirus (strain AD169)
C:Species: human cytomegalovirus, human herpesvirus 5
A:Note: host Homo sapiens (man)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 08-Oct-1999
C:Accession: S09777
R:Chen, M.S.; Bankier, A.T.; Beck, S.; Böhm, R.; Brown, C.M.; Cerny, R.; Horsnell, T
M.; Barrell, B.G.
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovi
A:Reference number: S09749; MUID:90269039
A:Accession: S09777
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-343 <CHE>
A:Cross-references: EMBL:X17403; NID:g959591; PIDN:CAA35447.1; PID:g959619
A:Note: this sequence was submitted to the EMBL Data Library, December 1989
C:Superfamily: human cytomegalovirus hypothetical protein ULI4
C:Keywords: glycoprotein; transmembrane protein
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-343/Product: hypothetical protein ULI4 #status predicted <MAT>
F:281-305/Domain: transmembrane #status predicted <TM>
F:160,255/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 2.2%; Score 8; DB 2; Length 343;
Best Local Similarity 100.0%; Pred. No. 13;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 240 AAGEASAV 247
 |||||
 DB 323 AAGEASAV 330

RESULT 12

C96680
 hypothetical protein F514.7 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: C96680
 R:Thelouis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 Nare, N.F.; Hughes, B.; Hulzer, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malt, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: C96680
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-398 <STO>
 A:Cross-References: GB:AE005173; NID:g2190544; PIDN:AA60908.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F514.7
 A:Map position: 1

Query Match 2.2%; Score 8; DB 2; Length 398;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 248 LAKAKAKA 255
 |||||
 DB 77 LAKAKAKA 84

RESULT 13

AG3259
 hypothetical protein BMEI0060 [imported] - Brucella melitensis (strain 16M)
 C:Species: Brucella melitensis
 C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
 C:Accession: AG3259
 R:Delvecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
 .; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess,
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
 A:Reference number: AD3252; PMID:11756688
 A:Accession: AG3259
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-400 <KUR>
 A:Cross-References: GB:AE008917; PIDN:AAL51242.1; PID:g17981931; GSPDB:GN00190
 A:Experimental source: strain 16M
 C:Genetics:
 A:Gene: BMEI0060
 A:Map position: 1

Query Match 2.2%; Score 8; DB 2; Length 400;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 323 PVPGRPDS 330
 |||||
 DB 21 PVPGRPDS 28

RESULT 14
 T18801
 hypothetical protein ZK131.11 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
 C:Accession: T18801; T27746
 R:Percy, C.
 submitted to the EMBL Data Library, March 1997
 A:Reference number: Z19024
 A:Accession: T18801
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-477 <WIL>
 A:Cross-References: EMBL:Z93373; PIDN:CAB07552.1; GSPDB:GN00020; CESP:ZK131.11
 A:Experimental source: clone C01B9
 R:Steward, C.
 submitted to the EMBL Data Library, December 1996
 A:Reference number: Z20413
 A:Accession: T27746
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-477 <WIL>
 A:Cross-References: EMBL:Z83245; PIDN:CAB05840.1; GSPDB:GN00020; CESP:ZK131.11
 A:Experimental source: clone ZK131
 C:Genetics:
 A:Gene: CESP:ZK131.11
 A:Map position: 2
 A:Intons: 20/2; 49/2; 113/1; 169/1; 260/2; 375/3; 432/3
 C:Superfamily: Caenorhabditis elegans hypothetical protein ZK131.11

Query Match 2.2%; Score 8; DB 2; Length 477;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 222 KQAOILAS 229
 |||||
 DB 457 KQAOILAS 464

RESULT 15

D98209
 aldehyde dehydrogenase dhds [imported] - Agrobacterium tumefaciens (strain C58, Cere
 C:Species: Agrobacterium tumefaciens
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
 C:Accession: D98209
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
 A.; Liu, F.; Woliam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz,
 Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
 A:Reference number: A97359; PMID:11743194
 A:Accession: D98209
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-493 <KUR>
 A:Cross-References: GB:AE007870; PIDN:AAK89198.1; PID:g15159016; GSPDB:GN00170
 C:Genetics:
 A:Gene: AGR_L1241
 A:Map position: linear chromosome
 C:Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology

Query Match 2.2%; Score 8; DB 2; Length 493;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 253 AKAEAIRI 260
 |||||
 DB 310 AKAEAIRI 317

Search completed: September 22, 2002, 18:32:59
Job time: 172 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 18:30:53 : Search time 39.01 Seconds
(without alignments)
2441.243 Million cell updates/sec

Title: US-09-898-216-1

Perfect score: 356

Sequence: 1 MLRAARGHMGPFAGELSTG.....RDVGTDA SXDELDYRKMS 356

Scoring table: OLIGO

Searched: 947138 seqs, 267508082 residues

Word size: 0

Total number of hits satisfying chosen parameters: 947138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: Pending_Patents_AA_New*

- 1: /cgn2-6/ptodata/2/paa/PCIT_NEW_COMB.pep.*
- 2: /cgn2-6/ptodata/2/paa/US06_NEW_COMB.pep.*
- 3: /cgn2-6/ptodata/2/paa/US07_NEW_COMB.pep.*
- 4: /cgn2-6/ptodata/2/paa/US08_NEW_COMB.pep.*
- 5: /cgn2-6/ptodata/2/paa/US09_NEW_COMB.pep.*
- 6: /cgn2-6/ptodata/2/paa/US10_NEW_COMB.pep.*
- 7: /cgn2-6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	354	99.4	356	US-09-898-216-1	Sequence 1, Appl.
2	218	61.2	356	US-09-629-469A-11010	Sequence 11010, A
3	218	61.2	356	US-10-219-793-150	Sequence 150, App
4	202	56.7	306	US-10-219-793-411	Sequence 411, App
5	202	56.7	356	US-60-389-987-1904	Sequence 1904, Ap
6	152	42.7	357	US-60-389-987-826	Sequence 826, App
7	143	40.2	342	US-60-389-987-768	Sequence 768, App
8	58	16.3	60	US-10-219-793-418	Sequence 418, App
9	58	16.3	60	US-10-219-793-416	Sequence 416, App
10	50	14.0	50	US-10-219-793-420	Sequence 420, App
11	45	12.6	45	US-10-219-793-422	Sequence 422, App
12	45	12.6	45	US-10-219-793-424	Sequence 424, App
13	35	9.8	46	US-10-219-793-426	Sequence 426, App
14	30	8.4	30	US-10-219-793-419	Sequence 419, App
15	28	7.9	28	US-10-219-793-417	Sequence 417, App
16	27	7.6	27	US-10-219-793-413	Sequence 413, App
17	26	7.3	26	US-10-219-793-412	Sequence 412, App
18	26	7.3	26	US-10-219-793-414	Sequence 414, App
19	25	7.0	25	US-10-219-793-421	Sequence 421, App
20	25	7.0	26	US-10-219-793-427	Sequence 427, App
21	24	6.7	24	US-10-219-793-423	Sequence 423, App
22	24	6.7	24	US-10-219-793-425	Sequence 425, App
23	24	6.7	29	US-10-219-793-415	Sequence 415, App
24	17	4.8	29	US-10-219-793-429	Sequence 429, App
25	17	4.8	30	US-10-219-793-428	Sequence 428, App
26	13	3.7	263	US-10-219-999-51693	Sequence 51693, A

27	13	3.7	275	6	US-10-219-999-39902	Sequence 39902, A
28	13	3.7	283	6	US-10-219-999-55341	Sequence 55341, A
29	10	2.8	10	6	US-10-227-616-58	Sequence 58, Appl
30	9	2.5	1109	6	US-10-179-131-6612	Sequence 6612, Ap
31	8	2.2	15	6	US-10-227-616-59	Sequence 59, Appl
32	8	2.2	43	6	US-10-182-995-20314	Sequence 20314, A
33	8	2.2	43	6	US-10-203-134-26567	Sequence 26567, A
34	8	2.2	43	6	US-10-203-136-26573	Sequence 26573, A
35	8	2.2	43	6	US-10-182-997-25978	Sequence 25978, A
36	8	2.2	43	6	US-10-182-997-18957	Sequence 18957, A
37	8	2.2	43	6	US-10-182-998-10609	Sequence 10609, A
38	8	2.2	43	6	US-10-203-135-25573	Sequence 25573, A
39	8	2.2	43	6	US-10-203-137-26806	Sequence 26806, A
40	8	2.2	43	6	US-10-203-138-10872	Sequence 10872, A
41	8	2.2	43	6	US-10-203-139-25711	Sequence 25711, A
42	8	2.2	187	6	US-10-219-999-49786	Sequence 49786, A
43	8	2.2	255	6	US-10-155-881-10456	Sequence 10456, A
44	8	2.2	471	7	US-60-360-036-11494	Sequence 11494, A
45	8	2.2	471	7	US-60-360-036-14424	Sequence 14424, A

ALIGNMENTS

RESULT 1
US-09-898-216-1
Sequence 1, Application US/09898216
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
GILL, Surya K.
TITLE OF INVENTION: NOVEL HUMAN MEMBRANE PROTEIN
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/898,216
FILING DATE: 02-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/781,562
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0181 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 356 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE: Consensus
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-898-216-1
Query Match 99.4%; Score 354; DB 5; Length 356;

Best Local Similarity 100.0%: Pred. No. 0;
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLARARGMGPFAGSLSTGFMPRSGRASSGLPRMTVVLFPVQOEAMVVERMGREHRIIE 60
Db 1 MLARARGMGPFAGSLSTGFMPRSGRASSGLPRMTVVLFPVQOEAMVVERMGREHRIIE 60
Qy 61 PGLNLLIPVLDRIIRYVQSLKEIIVINPEOSAVTLQIDGVLYLRIMDPYKASGYE 120
Db 61 PGLNLLIPVLDRIIRYVQSLKEIIVINPEOSAVTLQIDGVLYLRIMDPYKASGYE 120
Qy 121 DPEYAVTQAOITMRSELGKLSXDKVFRERESLNASTVDAINQAADCGICRLREIKKI 180
Db 121 DPEYAVTQAOITMRSELGKLSXDKVFRERESLNASTVDAINQAADCGICRLREIKKI 180
Qy 181 HYPVVKESMOMQVEARRKRAIVLESECTRESAINVAEGKQAOIILASEAKAEQINQA 240
Db 181 HYPVVKESMOMQVEARRKRAIVLESECTRESAINVAEGKQAOIILASEAKAEQINQA 240
Qy 241 AGEASAVLAKAKAKAEAIRILAAALTOHNGDAASLTVAEQYVSASFSLAKDSNTILLPS 300
Db 241 AGEASAVLAKAKAKAEAIRILAAALTOHNGDAASLTVAEQYVSASFSLAKDSNTILLPS 300
Qy 301 NPGDVTSMVAQAMGYGALTFRAPVGTTPDSLSSGSRDVOGTDAKXDELDRVKMS 356
Db 301 NPGDVTSMVAQAMGYGALTFRAPVGTTPDSLSSGSRDVOGTDAKXDELDRVKMS 356

RESULT 2

US-09-629-469A-11010
; Sequence 11010, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/09/629, 469A
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159, 590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183, 322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11010
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-629-469A-11010

Query Match 61.2%: Score 218; DB 5; Length 356;
Best Local Similarity 99.7%: Pred. No. 8e-197;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 27 RASSGLPRMTVVLFPVQOEAMVVERMGREHRIIEPGLNLLIPVLDRIIRYVQSLKEIIVN 86
Db 27 RASSGLPRMTVVLFPVQOEAMVVERMGREHRIIEPGLNLLIPVLDRIIRYVQSLKEIIVN 86
Qy 87 PQOSAVTLQIDGVLYLRIMDPYKASGYEDEPEYAVTQAOITMRSELGKLSXDKV 146
Db 87 PQOSAVTLQIDGVLYLRIMDPYKASGYEDEPEYAVTQAOITMRSELGKLSXDKV 146
Qy 147 FRERESLNASTVDAINQAADCGICRLREIKDIHYPVVKESMOMQVEARRKRAIVLE 206
Db 147 FRERESLNASTVDAINQAADCGICRLREIKDIHYPVVKESMOMQVEARRKRAIVLE 206
Qy 207 SECTRESAINVAEGKQAOIILASEAKAEQINQAAGEASAVLAKAKAEAIRILAAAL 266
Db 207 SECTRESAINVAEGKQAOIILASEAKAEQINQAAGEASAVLAKAKAEAIRILAAAL 266
Qy 267 QHNGDAASLTVAEQYVSASFSLAKDSNTILLPSNPGDVTSMVAQAMGYGALTFRAPVG 326
Db 267 QHNGDAASLTVAEQYVSASFSLAKDSNTILLPSNPGDVTSMVAQAMGYGALTFRAPVG 326
Qy 327 TPDSLSSGSRDVOGTDAK 345
Db 327 TPDSLSSGSRDVOGTDAK 345

RESULT 3

US-10-219-793-150
; Sequence 150, Application US/10219793
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; FILE REFERENCE: P2008P1C1
; CURRENT APPLICATION NUMBER: US/10/219, 793
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: 09/209, 462
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: PCT/US98/12125
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/049, 547
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049, 548
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049, 549
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049, 550
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049, 556
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049, 606
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049, 607
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049, 608
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049, 609
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049, 610
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049, 611
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/050, 901
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/052, 989
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/051, 919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055, 984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058, 665
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058, 668
; PRIOR FILING DATE: 1997-09-12


```

; PRIOR APPLICATION NUMBER: 60/058,669
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,750
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,971
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,972
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,975
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/060,834
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,841
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,844
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,865
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,059
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,060
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 737
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 150
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-219-793-150
```

```

Query Match          61.2%; Score 218; DB 6; Length 356;
Best Local Similarity 99.7%; Pred. No. 8e-197;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

QY 27 RASSGIPRTVYLVFVPOQEAHYVERMGRHRIIEPGLNIIIPVLDIRIVYQSLKEIYIN 86
DB 27 RASSGIPRTVYLVFVPOQEAHYVERMGRHRIIEPGLNIIIPVLDIRIVYQSLKEIYIN 86
QY 87 PEOSAVTLLNVTLOIDGVLYLRIMDPYKASVGEDEPEYAVTOLAORTMSELGKLSXDKY 146
DB 87 PEOSAVTLLNVTLOIDGVLYLRIMDPYKASVGEDEPEYAVTOLAORTMSELGKLSXDKY 146
QY 147 FRERESLNASIVDAINQADCMGICRCLRYEIKNDIHVPRVKESMOMQVEAERRKRATVLE 206
DB 147 FRERESLNASIVDAINQADCMGICRCLRYEIKNDIHVPRVKESMOMQVEAERRKRATVLE 206
QY 207 SESTRSAINVAEGKKQAOILASEAKAEQINQAGEASAVTLAKAKAKAEAIRILAAALT 266
DB 207 SESTRSAINVAEGKKQAOILASEAKAEQINQAGEASAVTLAKAKAKAEAIRILAAALT 266
QY 267 OHNGDAASLTVAEOYVSAFASKLAKDSNTLLPSPNGDVTSMVAOMGYGALTKAPVPG 326
DB 267 OHNGDAASLTVAEOYVSAFASKLAKDSNTLLPSPNGDVTSMVAOMGYGALTKAPVPG 326
QY 327 TPDSSLSSGSSRDVQGTAS 345
DB 327 TPDSSLSSGSSRDVQGTAS 345
```

```

RESULT 4
US-10-219-793-411
; Sequence 411, Application US/10219793
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; FILE REFERENCE: P2008PICI
; CURRENT APPLICATION NUMBER: US/10/219,793
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/209,462
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: PCT/US98/12125
; PRIOR FILING DATE: 1998-06-11
```

```

; PRIOR APPLICATION NUMBER: 60/049,547
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,548
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,549
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,550
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,556
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,606
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,607
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,608
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,609
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,610
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,611
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/050,901
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/052,989
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058,665
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,668
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,971
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,972
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,975
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/060,834
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,841
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,844
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,865
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,059
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,060
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 737
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 411
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-219-793-411
```

```

Query Match          56.7%; Score 202; DB 6; Length 306;
Best Local Similarity 100.0%; Pred. No. 8.4e-182;
Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 144 DKYFRERESLNASIVDAINQADCMGICRCLRYEIKNDIHVPRVKESMOMQVEAERRKRAT 203
DB 94 DKYFRERESLNASIVDAINQADCMGICRCLRYEIKNDIHVPRVKESMOMQVEAERRKRAT 153
```



```

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; FILE REFERENCE: P2008PICI
; CURRENT APPLICATION NUMBER: US/10/219,793
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/209,462
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: PCT/US98/12125
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/049,547
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,548
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,549
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,550
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,566
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,606
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,607
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,608
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,609
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,610
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,611
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/050,901
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/052,989
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058,665
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,668
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,669
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,750
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,971
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,972
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,975
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/060,834
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,841
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,844
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,865
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,059
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,060
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 737
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 416
; LENGTH: 60
; TYPE: prt
; ORGANISM: Homo sapiens
US-10-219-793-416

```

```

Query Match          16.3%: Score 58; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 9.9e-47;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 27 RASSGLPRNTVYLFVPEQEAAMVYERMGFRHRIIEPGINILIPVLDRIYVQSKEIYI 84
Db 3 RASSGLPRNTVYLFVPEQEAAMVYERMGFRHRIIEPGINILIPVLDRIYVQSKEIYI 60

RESULT 9
US-10-219-793-418
; Sequence 418, Application US/10219793
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; FILE REFERENCE: P2008PICI
; CURRENT APPLICATION NUMBER: US/10/219,793
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/209,462
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: PCT/US98/12125
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/049,547
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,548
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,549
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,550
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,566
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,606
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,607
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,608
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,609
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,610
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,611
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/050,901
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/052,989
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058,665
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,666
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,669
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,750
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,971
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,972
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,975
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/060,834
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,841
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,844

```

```

: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/060,865
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/061,059
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/061,060
: PRIOR FILING DATE: 1997-10-02
: NUMBER OF SEQ ID NOS: 737
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 418
: LENGTH: 60
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-219-793-418

Query Match          16.3%; Score 58; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 9.9e-47;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      85 NVPEGSATVLDNVTLQIDGVLYLRIMDPKASYGVEDPEYAVTOLAQTMRSELGKLS 142
DB      1 NVPEGSATVLDNVTLQIDGVLYLRIMDPKASYGVEDPEYAVTOLAQTMRSELGKLS 58

RESULT 10
US-10-219-793-420
: Sequence 420, Application US/10219793
: GENERAL INFORMATION:
: APPLICANT: Rosen et al
: TITLE OF INVENTION: 86 Human Secreted Proteins
: FILE REFERENCE: P2008P1C1
: CURRENT FILING DATE: 2002-08-16
: PRIOR APPLICATION NUMBER: 09/209,462
: PRIOR FILING DATE: 1998-12-11
: PRIOR APPLICATION NUMBER: PCT/US98/12125
: PRIOR FILING DATE: 1998-06-11
: PRIOR APPLICATION NUMBER: 60/049,547
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,548
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,549
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,550
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,566
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,606
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,607
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,608
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,609
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,610
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,611
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/050,901
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/052,989
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/051,919
: PRIOR FILING DATE: 1997-07-08
: PRIOR APPLICATION NUMBER: 60/055,984
: PRIOR FILING DATE: 1997-08-18
: PRIOR APPLICATION NUMBER: 60/058,665
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,668
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,669
```

```

: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,750
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,971
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,972
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,975
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/060,834
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/060,841
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/060,844
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/060,865
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/061,059
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/061,060
: NUMBER OF SEQ ID NOS: 737
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 420
: LENGTH: 50
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-219-793-420

Query Match          14.0%; Score 50; DB 6; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.9e-39;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      145 KVFREESINASTIVAINQADCGIRCRYEIKDTHVPPRKESKQMOV 194
DB      1 KVFREESINASTIVAINQADCGIRCRYEIKDTHVPPRKESKQMOV 50

RESULT 11
US-10-219-793-422
: Sequence 422, Application US/10219793
: GENERAL INFORMATION:
: APPLICANT: Rosen et al
: TITLE OF INVENTION: 86 Human Secreted Proteins
: FILE REFERENCE: P2008P1C1
: CURRENT FILING DATE: 2002-08-16
: PRIOR APPLICATION NUMBER: 09/209,462
: PRIOR FILING DATE: 1998-12-11
: PRIOR APPLICATION NUMBER: PCT/US98/12125
: PRIOR FILING DATE: 1998-06-11
: PRIOR APPLICATION NUMBER: 60/049,547
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,548
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,549
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,550
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,566
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,606
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,607
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,608
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,609
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,610
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,611
```

```
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/050,901
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/052,989
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058,665
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,668
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,669
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,750
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,971
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,972
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,975
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/060,834
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,841
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,844
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,865
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,059
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,060
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 737
; SOFTWARE: Patentln Ver. 2.0
; LENGTH: 45
; SEQ ID NO 422
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-219-793-422
```

```
Query Match          12.6%; Score 45; DB 6; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.4e-34;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 195 EAERRRRATVLESEGTRESAINVAEGKQAOILASEAEKAEQINQ 239
      ||||||||||||||||||||||||||||||||||||||||
Db 1 EAERRRRATVLESEGTRESAINVAEGKQAOILASEAEKAEQINQ 45
```

```
RESULT 12
US-10-219-793-424
; Sequence 424, Application US/10219793
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; FILE REFERENCE: P2008PICI
; CURRENT APPLICATION NUMBER: US/10/219,793
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/209,462
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: PCT/US98/12125
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/049,547
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,548
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,549
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,550
```

```
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,566
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,606
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,607
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,608
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,609
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,610
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,611
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/050,901
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/052,989
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058,665
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,668
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,669
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,750
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,971
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,972
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,975
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/060,834
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,841
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,844
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,865
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,059
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,060
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 737
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 424
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-219-793-424
```

```
Query Match          12.6%; Score 45; DB 6; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.4e-34;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 266 TOHNGDAASLTVAEOYVSASFSLAKDSNTILLPSPGQVTSNVA 310
      ||||||||||||||||||||||||||||||||||||||||
Db 1 TOHNGDAASLTVAEOYVSASFSLAKDSNTILLPSPGQVTSNVA 45
```

```
RESULT 13
US-10-219-793-426
; Sequence 426, Application US/10219793
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
```

```
FILE REFERENCE: P2008PICI
CURRENT APPLICATION NUMBER: US/10/219,793
CURRENT FILING DATE: 2002-08-16
PRIOR APPLICATION NUMBER: 09/209,462
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: PCT/US98/12125
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/049,547
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,548
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,549
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,550
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,566
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,606
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,607
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,608
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,609
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,610
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,611
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/050,901
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/052,989
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/051,919
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/055,984
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/058,665
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,668
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,669
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,750
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,971
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,972
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,975
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/060,834
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,841
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,844
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,865
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/061,059
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/061,060
PRIOR FILING DATE: 1997-10-02
NUMBER OF SEQ ID NOS: 737
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 426
LENGTH: 46
TYPE: PRT
ORGANISM: Homo sapiens
US-10-219-793-426
```

Query Match 9.8% Score 35; DB 6; Length 46;

```
Best Local Similarity 100.0%; Pred. No. 3,7e-25;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 311 QAMGYGALTAKAPVCPPTSLSSGSRDVGCTDAS 345
|||||
Db 1 QAMGYGALTAKAPVCPPTSLSSGSRDVGCTDAS 35

RESULT 14
US-10-219-793-419
Sequence 419, Application US/10219793
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 86 Human Secreted Proteins
FILE REFERENCE: P2008PICI
CURRENT APPLICATION NUMBER: US/10/219,793
CURRENT FILING DATE: 2002-08-16
PRIOR APPLICATION NUMBER: 09/209,462
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: PCT/US98/12125
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/049,547
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,548
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,549
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,550
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,566
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,606
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,607
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,608
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,609
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,610
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,611
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/050,901
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/052,989
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/051,919
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/055,984
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/058,665
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,668
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,669
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,750
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,971
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,972
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,975
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/060,834
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,841
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,844
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,865
PRIOR FILING DATE: 1997-10-02
```

```

; PRIOR APPLICATION NUMBER: 60/061,059
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,060
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 737
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 419
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-219-793-419

Query Match      8.4%; Score 30; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 1,3e-20;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 IDGVLYLRIMDPYKASYGVDEPEYAVTOLA 130
DB 1 IDGVLYLRIMDPYKASYGVDEPEYAVTOLA 30

RESULT 15
US-10-219-793-417
; Sequence 417, Application US/10219793
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; FILE REFERENCE: P2008P1
; CURRENT APPLICATION NUMBER: US/10/219,793
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/209,462
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: PCT/US98/12125
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/049,547
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,548
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,549
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,550
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,566
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,606
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,607
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,608
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,609
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,610
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,611
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/050,901
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/052,989
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058,665
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,668
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,669
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,750
; PRIOR FILING DATE: 1997-09-12

```

```

; PRIOR APPLICATION NUMBER: 60/058,971
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,972
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,975
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/060,834
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,841
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,844
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,865
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,059
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,060
; PRIOR FILING DATE: 1997-10-02
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 417
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-219-793-417

Query Match      7.9%; Score 28; DB 6; Length 28;
Best Local Similarity 100.0%; Pred. No. 9,4e-19;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 VPOQEAHVVERMGFRHRLIEPGLNIIIP 68
DB 1 VPOQEAHVVERMGFRHRLIEPGLNIIIP 28

```

Search completed: September 22, 2002, 18:35:49
Job time: 296 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 18:30:37 ; Search time 92.04 Seconds
(without alignments)
1361.420 Million cell updates/sec

Title: US-09-898-216-1

Perfect score: 356
Sequence: 1 MLARAARGHWGPFAGLSTG.....RDVQGTDSXDLELRVWMS 356

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 3502263 seqs, 351980561 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Pending_Patents_AA_Main: *
1: /cgn2_6/ptodata/1/paa/PCRTUS_COMB.pep: *
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep: *
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep: *
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep: *
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep: *
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep: *
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep: *
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep: *
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep: *
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep: *
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep: *
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep: *
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep: *
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep: *
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep: *
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep: *
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep: *
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep: *
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep: *
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep: *
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep: *
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep: *
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep: *
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep: *
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep: *
26: /cgn2_6/ptodata/1/paa/US60_COMB.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	218	61.2	356	1	PCT-US98-12125-145
2	218	61.2	356	16	US-09-209-462B-150
3	218	61.2	378	1	PCT-US01-03800A-2351
4	218	61.2	378	1	PCT-US01-04098A-3196
5	218	61.2	3007	1	PCT-US01-04942A-2222
6	218	61.2	3019	1	PCT-US01-08656-10594
7	202	56.7	305	16	US-09-298-733-36

8	202	56.7	305	16	US-09-298-733A-36	Sequence 36, App1
9	202	56.7	305	21	US-09-723-594-36	Sequence 36, App1
10	202	56.7	305	21	US-09-724-497-36	Sequence 36, App1
11	202	56.7	306	1	PCT-US98-12125-272	Sequence 272, App
12	202	56.7	306	16	US-09-209-462B-411	Sequence 411, App
13	198	55.6	358	1	PCT-US01-04098A-1228	Sequence 1228, App
14	58	16.3	60	16	US-09-209-462B-416	Sequence 416, App
15	58	16.3	60	16	US-09-209-462B-418	Sequence 418, App
16	51	14.3	51	26	US-60-163-233-2934	Sequence 2934, App
17	50	14.0	50	16	US-09-209-462B-420	Sequence 420, App
18	50	14.0	463	1	PCT-US01-04942A-598	Sequence 598, App
19	45	12.6	45	16	US-09-209-462B-422	Sequence 422, App
20	45	12.6	45	16	US-09-209-462B-424	Sequence 424, App
21	35	9.8	46	16	US-09-209-462B-426	Sequence 426, App
22	33	9.3	56	20	US-09-621-976-4903	Sequence 4903, App
23	33	9.3	56	26	US-60-147-499-4903	Sequence 4903, App
24	30	8.4	30	16	US-09-209-462B-419	Sequence 419, App
25	28	7.9	28	16	US-09-209-462B-417	Sequence 417, App
26	28	7.9	33	26	US-60-163-233-2731	Sequence 2731, App
27	28	7.9	33	26	US-60-164-764-1272	Sequence 1272, App
28	27	7.6	27	1	PCT-US98-12125-274	Sequence 274, App
29	27	7.6	27	16	US-09-209-462B-413	Sequence 413, App
30	26	7.3	26	1	PCT-US98-12125-273	Sequence 273, App
31	26	7.3	26	1	PCT-US98-12125-275	Sequence 275, App
32	26	7.3	26	16	US-09-209-462B-412	Sequence 412, App
33	26	7.3	26	16	US-09-209-462B-414	Sequence 414, App
34	25	7.0	25	16	US-09-209-462B-421	Sequence 421, App
35	25	7.0	26	16	US-09-209-462B-427	Sequence 427, App
36	24	6.7	24	16	US-09-209-462B-423	Sequence 423, App
37	24	6.7	24	16	US-09-209-462B-425	Sequence 425, App
38	24	6.7	70	1	PCT-US98-12125-276	Sequence 276, App
39	24	6.7	70	16	US-09-209-462B-415	Sequence 415, App
40	20	5.6	323	20	US-09-614-150-5415	Sequence 5415, App
41	20	5.6	323	26	US-60-191-637-5433	Sequence 5433, App
42	20	5.6	753	26	US-60-167-217-5529	Sequence 5529, App
43	17	4.8	39	16	US-09-209-462B-429	Sequence 429, App
44	17	4.8	30	16	US-09-209-462B-428	Sequence 428, App
45	13	3.7	265	21	US-09-708-427-57094	Sequence 57094, App

ALIGNMENTS

RESULT 1
PCT-US98-12125-145
Sequence 145, Application PC/TUS9812125
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 86 Human Secreted Proteins
NUMBER OF SEQUENCES: 318
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US98/12125
FILING DATE: Jan 01, 1990
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36, 373
REFERENCE/DOCKET NUMBER: P2008Complete

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 145:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 356 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; PCT-US98-12125-145

Query Match      61.2%; Score 218; DB 1; Length 356;
Best Local Similarity 99.7%; Pred. No. 4.6e-207;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 27 RASGGLPRNTVVLFPPOEAMVVERMGFRHRLLEPGNLILIPVLDRIITYVOSLKEIVIN 86
Db 27 RASGGLPRNTVVLFPPOEAMVVERMGFRHRLLEPGNLILIPVLDRIITYVOSLKEIVIN 86
Oy 87 PEOGAVTLDNVTLLQIDGVLRLIMDPYKASYGVEDPEYAVTQLAQTTMRSELGKLSXDKV 146
Db 87 PEOGAVTLDNVTLLQIDGVLRLIMDPYKASYGVEDPEYAVTQLAQTTMRSELGKLSXDKV 146
Oy 147 FRERESLNASTYVDALNQADCGICRCLREYETKDIHVPVRKESMOMOVEAEERRKATYLE 206
Db 147 FRERESLNASTYVDALNQADCGICRCLREYETKDIHVPVRKESMOMOVEAEERRKATYLE 206
Oy 207 SEGTRESAINVAEKKQKQIILASEEKAEOJNOAGASAVLAKAKAEAIRILAAALT 266
Db 207 SEGTRESAINVAEKKQKQIILASEEKAEOJNOAGASAVLAKAKAEAIRILAAALT 266
Oy 267 OHNGDAAASLTVAEQYSAFSLAKDSNTLLPSNPGDVTSMVAQAMGVGALTKAPVPG 326
Db 267 OHNGDAAASLTVAEQYSAFSLAKDSNTLLPSNPGDVTSMVAQAMGVGALTKAPVPG 326
Oy 327 TPDSLSSGSSRDVCGTDAS 345
Db 327 TPDSLSSGSSRDVCGTDAS 345

RESULT 2
US-09-209-462B-150
; Sequence 150, Application US/09209462B
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; FILE REFERENCE: P2008P1
; CURRENT APPLICATION NUMBER: US/09/209,462B
; CURRENT FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: PCT/US98/12125
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/049,547
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,548
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,549
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,550
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,566
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,606
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,607
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,608
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,609
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,610
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,611
; PRIOR FILING DATE: 1997-06-13
```

```

; PRIOR APPLICATION NUMBER: 60/050,901
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/052,989
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058,665
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,668
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,669
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,750
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,971
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,972
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,975
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/060,834
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,841
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,844
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,865
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,059
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,060
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 737
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 150
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-209-462B-150

Query Match      61.2%; Score 218; DB 16; Length 356;
Best Local Similarity 99.7%; Pred. No. 4.6e-207;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 27 RASGGLPRNTVVLFPPOEAMVVERMGFRHRLLEPGNLILIPVLDRIITYVOSLKEIVIN 86
Db 27 RASGGLPRNTVVLFPPOEAMVVERMGFRHRLLEPGNLILIPVLDRIITYVOSLKEIVIN 86
Oy 87 PEOGAVTLDNVTLLQIDGVLRLIMDPYKASYGVEDPEYAVTQLAQTTMRSELGKLSXDKV 146
Db 87 PEOGAVTLDNVTLLQIDGVLRLIMDPYKASYGVEDPEYAVTQLAQTTMRSELGKLSXDKV 146
Oy 147 FRERESLNASTYVDALNQADCGICRCLREYETKDIHVPVRKESMOMOVEAEERRKATYLE 206
Db 147 FRERESLNASTYVDALNQADCGICRCLREYETKDIHVPVRKESMOMOVEAEERRKATYLE 206
Oy 207 SEGTRESAINVAEKKQKQIILASEEKAEOJNOAGASAVLAKAKAEAIRILAAALT 266
Db 207 SEGTRESAINVAEKKQKQIILASEEKAEOJNOAGASAVLAKAKAEAIRILAAALT 266
Oy 267 OHNGDAAASLTVAEQYSAFSLAKDSNTLLPSNPGDVTSMVAQAMGVGALTKAPVPG 326
Db 267 OHNGDAAASLTVAEQYSAFSLAKDSNTLLPSNPGDVTSMVAQAMGVGALTKAPVPG 326
Oy 327 TPDSLSSGSSRDVCGTDAS 345
Db 327 TPDSLSSGSSRDVCGTDAS 345

RESULT 3
```

```
PCT-US01-03800A-2351
; Sequence 2351, Application PC/TUS0103800A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: PCT/US01/03800A
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 2351
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-03800A-2351
```

```
Query Match          61.2%; Score 218; DB 1; Length 378;
Best Local Similarity 99.7%; Pred. No. 4.9e-207;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 27 RASSGLPRNTVLFVPOQEAAMVERMGFRHRIIEPGINILIPVLDRIYVQSLKEIYIN 86
    |||
DB 49 RASSGLPRNTVLFVPOQEAAMVERMGFRHRIIEPGINILIPVLDRIYVQSLKEIYIN 108
    |||
QY 87 PEOASVTLDNVTIQLDGVLYLRIMDPYKASYGVDEPEYAVTQLAQTTRSELGKLSXDKV 146
    |||
DB 109 PEOASVTLDNVTIQLDGVLYLRIMDPYKASYGVDEPEYAVTQLAQTTRSELGKLSIDKV 168
    |||
QY 147 FRERESINASTIVAINQAADCGIRCLRYEIKDHPVPRKESMOMOVEAERRRRATVLE 206
    |||
DB 169 FRERESINASTIVAINQAADCGIRCLRYEIKDHPVPRKESMOMOVEAERRRRATVLE 228
    |||
QY 207 SEGTRESAINVAGGKKQAQILASEAEKAEQINOAGASAVLAKAKAKAIRILAAALT 266
    |||
DB 229 SEGTRESAINVAGGKKQAQILASEAEKAEQINOAGASAVLAKAKAKAIRILAAALT 288
    |||
QY 267 QHNGDAAASLTVAEQYVSASFSLAKDSNTILPSPNGDVTSMVAQAMGVGALTAKAPVG 326
    |||
DB 289 QHNGDAAASLTVAEQYVSASFSLAKDSNTILPSPNGDVTSMVAQAMGVGALTAKAPVG 348
    |||
QY 327 TPDSLSSGSSRDVOGTGDS 345
    |||
DB 349 TPDSLSSGSSRDVOGTGDS 367
    |||
```

```
RESULT 4
PCT-US01-04098A-3196
; Sequence 3196, Application PC/TUS0104098A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: PCT/US01/04098A
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
```

```
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 3196
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-04098A-3196
```

```
Query Match          61.2%; Score 218; DB 1; Length 378;
Best Local Similarity 99.7%; Pred. No. 4.9e-207;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 27 RASSGLPRNTVLFVPOQEAAMVERMGFRHRIIEPGINILIPVLDRIYVQSLKEIYIN 86
    |||
DB 49 RASSGLPRNTVLFVPOQEAAMVERMGFRHRIIEPGINILIPVLDRIYVQSLKEIYIN 108
    |||
QY 87 PEOASVTLDNVTIQLDGVLYLRIMDPYKASYGVDEPEYAVTQLAQTTRSELGKLSXDKV 146
    |||
DB 109 PEOASVTLDNVTIQLDGVLYLRIMDPYKASYGVDEPEYAVTQLAQTTRSELGKLSIDKV 168
    |||
QY 147 FRERESINASTIVAINQAADCGIRCLRYEIKDHPVPRKESMOMOVEAERRRRATVLE 206
    |||
DB 169 FRERESINASTIVAINQAADCGIRCLRYEIKDHPVPRKESMOMOVEAERRRRATVLE 228
    |||
QY 207 SEGTRESAINVAGGKKQAQILASEAEKAEQINOAGASAVLAKAKAKAIRILAAALT 266
    |||
DB 229 SEGTRESAINVAGGKKQAQILASEAEKAEQINOAGASAVLAKAKAKAIRILAAALT 288
    |||
QY 267 QHNGDAAASLTVAEQYVSASFSLAKDSNTILPSPNGDVTSMVAQAMGVGALTAKAPVG 326
    |||
DB 289 QHNGDAAASLTVAEQYVSASFSLAKDSNTILPSPNGDVTSMVAQAMGVGALTAKAPVG 348
    |||
QY 327 TPDSLSSGSSRDVOGTGDS 345
    |||
DB 349 TPDSLSSGSSRDVOGTGDS 367
    |||
```

```
RESULT 5
PCT-US01-04942A-222
; Sequence 222, Application PC/TUS0104942A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: PCT/US01/04942A
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 222
; LENGTH: 3007
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-04942A-222
```

```
Query Match          61.2%; Score 218; DB 1; Length 3007;
```

Best Local Similarity 99.7%, Pred. No. 3.8e-206;
Matches 318: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Oy 27 RASSGLPNTVLEFVPOQEAAMVERMGRFHRIIEPGLNIIIPVLDRIIRYQSLKEIVIN 86
Db 27 RASSGLPNTVLEFVPOQEAAMVERMGRFHRIIEPGLNIIIPVLDRIIRYQSLKEIVIN 86
Oy 87 PEQSAVTLDNVTLQIDGVLYLRIMDPYKASYGVEDPEYAVTQLAQTTMRSELGKLSXKV 146
Db 87 PEQSAVTLDNVTLQIDGVLYLRIMDPYKASYGVEDPEYAVTQLAQTTMRSELGKLSXKV 146
Oy 147 FRERESLNASIVDAINQAADCGIRCLREYKIDIHVPPRVKESMOMOVEAERKRATYLE 206
Db 147 FRERESLNASIVDAINQAADCGIRCLREYKIDIHVPPRVKESMOMOVEAERKRATYLE 206
Oy 207 SEGTRESAINVAEGKKQAOIILASEEKAEOINOAGASAVLAKAKAEAIRILAAALT 266
Db 207 SEGTRESAINVAEGKKQAOIILASEEKAEOINOAGASAVLAKAKAEAIRILAAALT 266
Oy 267 QHNGDAASLTVAEOYVSFAFSKLAKDSNTIILPSNPGDVTSMVAQAMGYGALTAPVPG 326
Db 267 QHNGDAASLTVAEOYVSFAFSKLAKDSNTIILPSNPGDVTSMVAQAMGYGALTAPVPG 326
Oy 327 TPDSLSGSSSRDVOGTDS 345
Db 327 TPDSLSGSSSRDVOGTDS 345
```

RESULT 6
PCT-US01-08656-10594

; Sequence 10594, Application PC/TUS0108656
; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066

; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: 09/522,929

; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160

; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994

; SOFTWARE: Custom

; SEQ ID NO 10594

; LENGTH: 3019

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: DOMAIN

; LOCATION: (115)..(144)

; OTHER INFORMATION: Band 7 protein family proteins domain identified by EMATRIX.

; OTHER INFORMATION: accession number BL01270C, p-value=6.745e-17, raw score of 16.91

; NAME/KEY: misc_feature

; LOCATION: (1)..(3019)

; OTHER INFORMATION: Xaa - X or * as defined in Table 2

PCT-US01-08656-10594

Query Match 61.2%; Score 218; DB 1; Length 3019;

Best Local Similarity 99.7%; Pred. No. 3.8e-206;
Matches 318: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Oy 27 RASSGLPNTVLEFVPOQEAAMVERMGRFHRIIEPGLNIIIPVLDRIIRYQSLKEIVIN 86
Db 27 RASSGLPNTVLEFVPOQEAAMVERMGRFHRIIEPGLNIIIPVLDRIIRYQSLKEIVIN 86
Oy 87 PEQSAVTLDNVTLQIDGVLYLRIMDPYKASYGVEDPEYAVTQLAQTTMRSELGKLSXKV 146
Db 87 PEQSAVTLDNVTLQIDGVLYLRIMDPYKASYGVEDPEYAVTQLAQTTMRSELGKLSXKV 146
```

```
Db 87 PEQSAVTLDNVTLQIDGVLYLRIMDPYKASYGVEDPEYAVTQLAQTTMRSELGKLSXKV 146
Oy 147 FRERESLNASIVDAINQAADCGIRCLREYKIDIHVPPRVKESMOMOVEAERKRATYLE 206
Db 147 FRERESLNASIVDAINQAADCGIRCLREYKIDIHVPPRVKESMOMOVEAERKRATYLE 206
Oy 207 SEGTRESAINVAEGKKQAOIILASEEKAEOINOAGASAVLAKAKAEAIRILAAALT 266
Db 207 SEGTRESAINVAEGKKQAOIILASEEKAEOINOAGASAVLAKAKAEAIRILAAALT 266
Oy 267 QHNGDAASLTVAEOYVSFAFSKLAKDSNTIILPSNPGDVTSMVAQAMGYGALTAPVPG 326
Db 267 QHNGDAASLTVAEOYVSFAFSKLAKDSNTIILPSNPGDVTSMVAQAMGYGALTAPVPG 326
Oy 327 TPDSLSGSSSRDVOGTDS 345
Db 327 TPDSLSGSSSRDVOGTDS 345
```

RESULT 7

US-09-298-733-36
; Sequence 36, Application US/09298733
; GENERAL INFORMATION:

; APPLICANT: Valenzuela, Dario

; APPLICANT: Yuan, Olive

; APPLICANT: Hoffman, Heidi

; APPLICANT: Hall, Jeff

; APPLICANT: Raplejo, Peter

; APPLICANT: Adams, R. Mark

; APPLICANT: Agostino, Michael J.

; APPLICANT: Rechel, Kim

; APPLICANT: Howes, Steve H.

; TITLE OF INVENTION: Secreted proteins and polynucleotides

; FILE REFERENCE: AG199-02PM

; CURRENT FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: US/09/298,733

; PRIOR FILING DATE: 1998-04-24
; PRIOR APPLICATION NUMBER: 60/082,961

; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086,402

; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088,994

; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089,163

; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091,619

; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 36

; LENGTH: 305

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-298-733-36

Query Match 56.7%; Score 202; DB 16; Length 305;

Best Local Similarity 100.0%; Pred. No. 3e-191;
Matches 202: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Oy 144 DKYFRERESLNASIVDAINQAADCGIRCLREYKIDIHVPPRVKESMOMOVEAERKRAT 203
Db 93 DKYFRERESLNASIVDAINQAADCGIRCLREYKIDIHVPPRVKESMOMOVEAERKRAT 152
Oy 204 VLESEGTRESAINVAEGKKQAOIILASEEKAEOINOAGASAVLAKAKAEAIRILAA 263
Db 153 VLESEGTRESAINVAEGKKQAOIILASEEKAEOINOAGASAVLAKAKAEAIRILAA 212
Oy 264 ALTOHNGDAASLTVAEOYVSFAFSKLAKDSNTIILPSNPGDVTSMVAQAMGYGALTAP 323
Db 213 ALTOHNGDAASLTVAEOYVSFAFSKLAKDSNTIILPSNPGDVTSMVAQAMGYGALTAP 272
Oy 324 VPSTPDSLSGSSSRDVOGTDS 345
```


US-09-724-497-36

Query Match	Score 202;	DB 21;	Length 305;
56.78;			
100.00			

```
Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

```
Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

Qy	144	DKVVERKESLVAASLVDAIINQAADCMGJICLCTEYETKDIHVHPRAESQOMQVEAEERRRAT	203
Db	93	DKVVERESLVAASLVDAIINQAADCMGJICLCTEYETKDIHVHPRAESQOMQVEAEERRRAT	152
Qy	204	VLESEBGTRESAINVAEKKKQAOILLASEEKAKEQINOQAAGEASVLAKAKKAEAIRILAA	263
Db	153	VLESEBGTRESAINVAEKKKQAOILLASEEKAKEQINOQAAGEASVLAKAKKAEAIRILAA	212
Qy	264	ALTOHNGDAASLTVABEYVSASFELKLAKDSNTILLPSNPEDVYSMAVQAAMGVGALTAP	323
Db	213	ALTOHNGDAASLTVABEYVSASFELKLAKDSNTILLPSNPEDVYSMAVQAAMGVGALTAP	272
Qy	324	VPGRPDSLSSGSSRDVQGTDAAS	345
Db	273	VPGRPDSLSSGSSRDVQGTDAAS	294

```

RESULT 11
PCT-US98-12125-272
; Sequence 272, Application PC/TUS9812125
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; NUMBER OF SEQUENCES: 318
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; Zip: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 Inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/12125
; FILING DATE: Jan 01, 1900
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brooks
; REGISTRATION NUMBER: 36, 373
; REFERENCE/DOCKET NUMBER: P2008complete
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 272:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 306 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; PCT-US98-12125-272

```

Query Match	56.7%	Score 202:	DB 1:	Length 306;
Best Local Similarity	100.0%	Pred. No.	3e-191;	
Matches 202:	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

Oy	144 DKYFRERESLNASTIDAINQADCMGIRCLRYEIKDHPVPRVESMOMOVEAERRKAT	203
D6	94 DKYFRERESLNASTIDAINQADCMGIRCLRYEIKDHPVPRVESMOMOVEAERRKAT	153

QY	204	VLESEGTRESAINNAEKKQAOIILASEEKAEQINOAGAEGSAYLAKAKKAEIRTLAA	263
Db	154	VLESEGTRESAINNAEKKQAOIILASEEKAEQINOAGAEGSAYLAKAKKAEIRTLAA	213
QY	264	ALTOHNGDAASLTVAEDVYSAFSKLAKDSNTILLPSNPGDVTSMAVQAANGVYCALTKAP	323
Db	214	ALTOHNGDAASLTVAEDVYSAFSKLAKDSNTILLPSNPGDVTSMAVQAANGVYCALTKAP	273
QY	324	VPGTDSLSSGSRDVOGTDAS	345
Db	274	VPGTDSLSSGSRDVOGTDAS	295

```

US-09-209-462B-411
: Sequence 411, Application US/09209462B
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: 86 Human Secreted Proteins
: FILE REFERENCE: P2006P1
: CURRENT APPLICATION NUMBER: US/09/209,462B
: CURRENT FILING DATE: 1998-12-11
: PRIOR APPLICATION NUMBER: PCT/US98/12125
: PRIOR FILING DATE: 1998-06-11
: PRIOR APPLICATION NUMBER: 60/049,547
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,548
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,549
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,550
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,566
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,606
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,607
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,608
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,609
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,610
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/049,611
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/050,901
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/052,989
: PRIOR FILING DATE: 1997-06-13
: PRIOR APPLICATION NUMBER: 60/051,919
: PRIOR FILING DATE: 1997-07-08
: PRIOR APPLICATION NUMBER: 60/055,984
: PRIOR FILING DATE: 1997-08-18
: PRIOR APPLICATION NUMBER: 60/058,665
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,668
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,669
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,750
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,971
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,972
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/058,975
: PRIOR FILING DATE: 1997-09-12
: PRIOR APPLICATION NUMBER: 60/060,834
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/060,841
: PRIOR FILING DATE: 1997-10-02
: PRIOR APPLICATION NUMBER: 60/060,844

```

```

; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,865
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,059
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,060
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 737
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 411
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-209-462B-411
```

```
Query Match          56.7%  Score 202:  DB 16:  Length 306:
Best Local Similarity 100.0%  Pred. No. 3e-191:
Matches 202:  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

```
QY 144 DKYFRRESLNASIVDAINQADCMGICRCLRYEIKDIHPPRVKESMOMQVEAERRRRAT 203
    |||||||
DB 94 DKYFRRESLNASIVDAINQADCMGICRCLRYEIKDIHPPRVKESMOMQVEAERRRRAT 153
    |||||||
QY 204 VLESEGTRESAINVAEGKKQAOQLLASEAEKAEQINQAGEASAVLAKAKAKAEAIRILAA 263
    |||||||
DB 154 VLESEGTRESAINVAEGKKQAOQLLASEAEKAEQINQAGEASAVLAKAKAKAEAIRILAA 213
    |||||||
QY 264 ALTOHNGDAAASLTVAEOYVSASFSLAKDSNTLLPSNPGDVTSMVAQAMGVGALTAKAP 323
    |||||||
DB 214 ALTOHNGDAAASLTVAEOYVSASFSLAKDSNTLLPSNPGDVTSMVAQAMGVGALTAKAP 273
    |||||||
QY 324 VPGTPDLSLSSGSSRDYQGTDAAS 345
    |||||||
DB 274 VPGTPDLSLSSGSSRDYQGTDAAS 295
    |||||||
```

```
RESULT 13
PCT-US01-04098A-1228
; Sequence 1228, Application PC/TUS0104098A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: PCT/US01/04098A
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1228
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-04098A-1228
```

```
Query Match          55.6%  Score 198:  DB 1:  Length 358:
Best Local Similarity 99.7%  Pred. No. 3.3e-187:
Matches 298:  Conservative 0;  Mismatches 1;  Indels 0;  Gaps 0;
```

```
QY 27 RASSGLPRNTVLFVPOQEAQVVERKGRFRRLIEPGINILIPVLDRIYQSLKEIYINV 86
    |||||||
DB 27 RASSGLPRNTVLFVPOQEAQVVERKGRFRRLIEPGINILIPVLDRIYQSLKEIYINV 86
    |||||||
QY 87 PEOSAVTLDNVTLQIDGVLYLRTIMDPYKASGVDEPEYATQLAQTTMRSELGLSKSDKY 146
    |||||||
DB 87 PEOSAVTLDNVTLQIDGVLYLRTIMDPYKASGVDEPEYATQLAQTTMRSELGLSKSDKY 146
    |||||||
QY 147 FRERESLNASIVDAINQADCMGICRCLRYEIKDIHPPRVKESMOMQVEAERRRRATVLE 206
    |||||||
DB 147 FRERESLNASIVDAINQADCMGICRCLRYEIKDIHPPRVKESMOMQVEAERRRRATVLE 206
    |||||||
QY 207 SEGTRESAINVAEGKKQAOQLLASEAEKAEQINQAGEASAVLAKAKAKAEAIRILAAALT 266
    |||||||
DB 207 SEGTRESAINVAEGKKQAOQLLASEAEKAEQINQAGEASAVLAKAKAKAEAIRILAAALT 266
    |||||||
QY 267 QHNGDAAASLTVAEOYVSASFSLAKDSNTLLPSNPGDVTSMVAQAMGVGALTAKAPV 325
    |||||||
DB 267 QHNGDAAASLTVAEOYVSASFSLAKDSNTLLPSNPGDVTSMVAQAMGVGALTAKAPV 325
    |||||||
```

```
RESULT 14
US-09-209-462B-416
; Sequence 416, Application US/09209462B
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; FILE REFERENCE: P2008P1
; CURRENT APPLICATION NUMBER: US/09/209,462B
; CURRENT FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: PCT/US98/12125
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/049,547
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,548
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,549
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,550
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,566
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,606
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,607
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,608
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,609
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,610
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,611
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/050,901
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/052,989
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058,665
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,668
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,669
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,750
```

```

; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058, 971
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058, 972
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058, 975
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/060, 834
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060, 841
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060, 844
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060, 865
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061, 059
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061, 060
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 737
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 416
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-209-462B-416
```

```

Query Match      16.3%; Score 58; DB 16; Length 60;
Best Local Similarity 100.0%; Pred. No. 6e-49;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 27 RASGLPNTVVLVPPQGEAMVVERMGRFHRILEPGLNLIIPVLDRIYVOSLKEIIV 84
      |||||||
DB 3 RASGLPNTVVLVPPQGEAMVVERMGRFHRILEPGLNLIIPVLDRIYVOSLKEIIV 60
```

```

RESULT 15
US-09-209-462B-416
; Sequence 418, Application US/09209462B
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; FILE REFERENCE: P2008P1
; CURRENT APPLICATION NUMBER: US/09/209,462B
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: PCT/US98/12125
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/049, 547
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049, 548
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049, 549
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049, 550
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049, 566
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049, 606
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049, 607
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049, 608
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049, 609
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049, 610
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049, 611
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/050, 901
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/052, 989
```

```

; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/051, 919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055, 984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058, 665
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058, 668
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058, 669
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058, 750
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058, 971
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058, 972
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058, 975
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/060, 834
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060, 841
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060, 844
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060, 865
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061, 059
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061, 060
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 737
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 418
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-209-462B-416
```

```

Query Match      16.3%; Score 58; DB 16; Length 60;
Best Local Similarity 100.0%; Pred. No. 6e-49;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 85 NVPEQSATVLDNVTLQIDGVLYLRIMDPYKASYGDEPPEYATOLAQTMRSELKLS 142
      |||||||
DB 1 NVPEQSATVLDNVTLQIDGVLYLRIMDPYKASYGDEPPEYATOLAQTMRSELKLS 58
```

Search completed: September 22, 2002, 18:34:49
Job time: 252 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 18:26:32 ; Search time 15.91 Seconds
(without alignments)
546.544 Million cell updates/sec

Title: US-09-898-216-1

Perfect score: 356

Sequence: 1 MLARAARGHMGPFAGELSTG.....RDVQGTDSXDELDPRVKMS 356

Scoring table:

OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTC05.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	354	99.4	356	1	US-08-781-562-1
2	7	2.0	16	1	US-08-346-849-25
3	7	2.0	16	1	US-08-346-849-32
4	7	2.0	16	2	US-08-293-284A-25
5	7	2.0	16	2	US-08-293-284A-32
6	7	2.0	24	4	US-08-993-008A-4
7	7	2.0	29	1	US-08-152-486-3
8	7	2.0	29	1	US-08-303-025-3
9	7	2.0	29	2	US-08-677-304-3
10	7	2.0	35	4	US-08-436-703B-8
11	7	2.0	48	4	US-08-995-172-18
12	7	2.0	48	4	US-08-993-008A-5
13	7	2.0	56	4	US-08-993-008A-6
14	7	2.0	60	2	US-08-769-211-2
15	7	2.0	61	4	US-08-995-172-6
16	7	2.0	61	4	US-08-995-172-7
17	7	2.0	100	2	US-08-460-890A-64
18	7	2.0	100	3	US-08-167-641C-64
19	7	2.0	100	4	US-08-460-971A-64
20	7	2.0	100	4	US-08-462-040-64
21	7	2.0	283	4	US-09-036-987A-13
22	7	2.0	283	4	US-09-370-700-13
23	7	2.0	364	4	US-08-891-322-6
24	7	2.0	364	4	US-09-277-019-6
25	7	2.0	381	1	US-08-781-562-6
26	7	2.0	422	2	US-09-067-351-3
27	7	2.0	422	4	US-09-360-490-3

28	7	2.0	432	3	US-09-075-087-2	Sequence 2, Appl1
29	7	2.0	432	4	US-09-472-971-1	Sequence 1, Appl1
30	7	2.0	433	1	US-08-346-849-2	Sequence 2, Appl1
31	7	2.0	433	2	US-08-293-284A-2	Sequence 2, Appl1
32	7	2.0	553	3	US-09-083-351-2	Sequence 2, Appl1
33	7	2.0	553	4	US-09-083-352-2	Sequence 2, Appl1
34	7	2.0	684	4	US-08-965-762-2	Sequence 2, Appl1
35	7	2.0	693	2	US-08-380-403A-2	Sequence 2, Appl1
36	7	2.0	693	2	US-08-380-403A-5	Sequence 5, Appl1
37	7	2.0	693	2	US-08-895-628-2	Sequence 5, Appl1
38	7	2.0	693	2	US-08-895-628-5	Sequence 5, Appl1
39	7	2.0	713	3	US-09-335-409-11	Sequence 11, Appl1
40	7	2.0	713	4	US-09-568-102-11	Sequence 11, Appl1
41	7	2.0	713	4	US-09-567-969-11	Sequence 11, Appl1
42	7	2.0	713	4	US-09-568-480-11	Sequence 11, Appl1
43	7	2.0	713	4	US-09-568-486-11	Sequence 11, Appl1
44	7	2.0	713	4	US-09-568-472-11	Sequence 11, Appl1
45	7	2.0	740	3	US-08-891-322-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-08-781-562-1
Sequence 1, Application US/08781562
Patent No. 5763589
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: NOVEL HUMAN MEMBRANE PROTEIN
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,562
FILING DATE: Herewith
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0181 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 356 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE: Consensus
US-08-781-562-1
Query Match 99.4%; Score 354; DB 1; Length 356;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLARAARHGMPFAGLSTGFWPRSGRASSGLPRNTVLEFVPOEAWYVERMGFRHILE 60
Db 1 MLARAARHGMPFAGLSTGFWPRSGRASSGLPRNTVLEFVPOEAWYVERMGFRHILE 60

Qy 61 PCLNTLLIPVLDRIYVOSLKEIVINWPEOSAVTLDNVTLDIGVLYLRIMDPYKASGYVE 120
Db 61 PCLNTLLIPVLDRIYVOSLKEIVINWPEOSAVTLDNVTLDIGVLYLRIMDPYKASGYVE 120

Qy 121 DPEYAVTOLAQTMRSEGLKSLXDKVFRERESLNASIVDAINQADCMGIRCLREYIKDI 180
Db 121 DPEYAVTOLAQTMRSEGLKSLXDKVFRERESLNASIVDAINQADCMGIRCLREYIKDI 180

Qy 181 HVPPEKESMQVQVBAERKRKATVLESECTRESAINVAEGKQAOILASEAKAQINDA 240
Db 181 HVPPEKESMQVQVBAERKRKATVLESECTRESAINVAEGKQAOILASEAKAQINDA 240

Qy 241 AGEASAVLAKAKAKAARILAAALTOHNGDAASLVAEOYVSAFSLAKDSNTLLPS 300
Db 241 AGEASAVLAKAKAKAARILAAALTOHNGDAASLVAEOYVSAFSLAKDSNTLLPS 300

Qy 301 NPGDVTSMVAQAMGYGALTKAPVGTPTDLSGSSSRDVOGTDAKDDELDRVKMS 356
Db 301 NPGDVTSMVAQAMGYGALTKAPVGTPTDLSGSSSRDVOGTDAKDDELDRVKMS 356

RESULT 2
US-08-346-849-25
; Sequence 25, Application US/08346849
; Patent No. 5670483
; GENERAL INFORMATION:
; APPLICANT: Zhang, Shuguang
; APPLICANT: Lockshin, Curtis
; APPLICANT: Rich, Alexander
; APPLICANT: Holmes, Todd
; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
; NUMBER OF SEQUENCES: 64
; CURRENT APPLICATION DATA:
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Millitia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/346,849
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,326
; FILING DATE: 28 DECEMBER 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: MIT-6008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-346-849-25

Query Match 2.0%; Score 7; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 249 AKAKAA 255
Db 2 AKAKAA 8

RESULT 3
US-08-346-849-32
; Sequence 32, Application US/08346849
; Patent No. 5670483
; GENERAL INFORMATION:
; APPLICANT: Zhang, Shuguang
; APPLICANT: Lockshin, Curtis
; APPLICANT: Rich, Alexander
; APPLICANT: Holmes, Todd
; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
; NUMBER OF SEQUENCES: 64
; CURRENT APPLICATION DATA:
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Millitia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/346,849
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,326
; FILING DATE: 28 DECEMBER 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: MIT-6008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-346-849-32

Query Match 2.0%; Score 7; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 251 AKAKAA 257
Db 1 AKAKAA 7

RESULT 4
US-08-293-284A-25
; Sequence 25, Application US/08293284A
; Patent No. 5955343
; GENERAL INFORMATION:

```

: APPLICANT: Holmes, Todd
: APPLICANT: Zhang, Shuangang
: APPLICANT: Rich, Alexander
: APPLICANT: DiPersio, C. Michael
: APPLICANT: Lockshin, Curtis
: TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
: TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
: TITLE OF INVENTION: THEREFOR
: NUMBER OF SEQUENCES: 64
: CORRESPONDENCE ADDRESS:
: ADDRESS: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
: STREET: Two Millitia Drive
: CITY: Lexington
: STATE: Massachusetts
: COUNTRY: U.S.A.
: ZIP: 02173-4799
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: PC-DOS/MS-DOS
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/293,284A
: FILING DATE: 22-AUG-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/973,326
: FILING DATE: 28-DEC-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Brook, David E.
: REGISTRATION NUMBER: 22,592
: REFERENCE/DOCKET NUMBER: MIT-6008A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 861-6240
: TELEFAX: (617) 861-9540
: INFORMATION FOR SEQ ID NO: 25:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 16 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-08-293-284A-25

```

```

Query Match          2.0%; Score 7; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 249 AKAKAKA 255
 Db 2 AKAKAKA 8

```

RESULT 5
: US-08-293-284A-32
: Sequence 32, Application US/08293284A
: Patent No. 5955343
: GENERAL INFORMATION:
: APPLICANT: Holmes, Todd
: APPLICANT: Zhang, Shuangang
: APPLICANT: Rich, Alexander
: APPLICANT: DiPersio, C. Michael
: APPLICANT: Lockshin, Curtis
: TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
: TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
: TITLE OF INVENTION: THEREFOR
: NUMBER OF SEQUENCES: 64
: CORRESPONDENCE ADDRESS:
: ADDRESS: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
: STREET: Two Millitia Drive
: CITY: Lexington
: STATE: Massachusetts
: COUNTRY: U.S.A.

```

```

: ZIP: 02173-4799
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: PC-DOS/MS-DOS
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/293,284A
: FILING DATE: 22-AUG-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/973,326
: FILING DATE: 28-DEC-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Brook, David E.
: REGISTRATION NUMBER: 22,592
: REFERENCE/DOCKET NUMBER: MIT-6008A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 861-6240
: TELEFAX: (617) 861-9540
: INFORMATION FOR SEQ ID NO: 32:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 16 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-08-293-284A-32

```

```

Query Match          2.0%; Score 7; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 251 AKAKAKA 257
 Db 1 AKAKAKA 7

```

RESULT 6
: US-08-993-008A-4
: Sequence 4, Application US/08993008A
: Patent No. 6153596
: GENERAL INFORMATION:
: APPLICANT: Liotta, Dennis C.
: APPLICANT: Petros, John A.
: APPLICANT: Wei, Shioh-Jyi
: APPLICANT: Kair, Joan F.
: APPLICANT: Pohl, Jan
: TITLE OF INVENTION: Polycationic Oligomers -
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESS: Greenlee, Winner and Sullivan
: STREET: 5370 Manhattan Circle, Suite 201
: CITY: Boulder
: STATE: CO
: COUNTRY: US
: ZIP: 80303
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: PC-DOS/MS-DOS
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/993,008A
: FILING DATE: 18-DEC-1997
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/032,436
: FILING DATE: 18-DEC-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Sullivan, Sally A.
: REGISTRATION NUMBER: 32,064
: REFERENCE/DOCKET NUMBER: 33-95

```

TELECOMMUNICATION INFORMATION:
TELEPHONE: 303-499-8080
TELEFAX: 303-499-8089
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-993-008A-4

Query Match 2.0%; Score 7; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 8.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 249 AKAKAKA 255
|||||
Db 2 AKAKAKA 8

RESULT 7
US-08-152-488-3
Sequence 3, Application US/08152488
Patent No. 5534619
GENERAL INFORMATION:
APPLICANT: Wakefield, Thomas W.
APPLICANT: Andrews, Philip C.
TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
TITLE OF INVENTION: ANTICOAGULATION REVERSAL
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Benita J. Rohm, Esq.
STREET: 512 Springfield Avenue
CITY: Cranford
STATE: New Jersey
COUNTRY: United States of America
ZIP: 07016-1811
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 6; ASCII (DOS)Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/152,488
FILING DATE: 12-NOV-1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/08069
FILING DATE: 14-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Rohm, Benita J.
REGISTRATION NUMBER: 28,664
REFERENCE/DOCKET NUMBER: RM-7MG
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-276-3344
TELEFAX: 908-276-5543
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: N/A
PUBLICATION INFORMATION:
AUTHORS: N/A

TITLE: N/A
PUBLICATION INFORMATION:
DOCUMENT NUMBER: PCT/US92/08069
FILING DATE: 14-AUG-1993
US-08-152-488-3

Query Match 2.0%; Score 7; DB 1; Length 29;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 249 AKAKAKA 255
|||||
Db 4 AKAKAKA 10

RESULT 8
US-08-303-025-3
Sequence 3, Application US/08303025
Patent No. 5614494
GENERAL INFORMATION:
APPLICANT: Wakefield, Thomas W.
APPLICANT: Andrews, Philip C.
APPLICANT: Stanley, James C.
TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
TITLE OF INVENTION: ANTICOAGULATION REVERSAL
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Benita J. Rohm, Esq.
STREET: 150 West Jefferson, Suite 2500
CITY: Detroit
STATE: Michigan
COUNTRY: United States of America
ZIP: 48226-4415
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy diskette 3.5" 1.44MB
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS V.6.22
SOFTWARE: Wordperfect 6.1; ASCII (DOS)Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/303,025
FILING DATE: 08-SEPT-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/06829
FILING DATE: 14-AUG-1992
APPLICATION NUMBER: US 08/152,488
FILING DATE: 12-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Rohm, Benita J.
REFERENCE/DOCKET NUMBER: 7MH-060548-00231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 313-496-7622
TELEFAX: 313-496-8454
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: N/A
PUBLICATION INFORMATION:
AUTHORS: N/A
TITLE: N/A
DOCUMENT NUMBER: PCT/US92/08069
FILING DATE: 14-AUG-1993
US-08-303-025-3

Query Match 2.0%; Score 7; DB 1; Length 29;

Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 AKAKAKA 255
DB 4 AKAKAKA 10

RESULT 9
US-08-677-304-3
; Sequence 3, Application US/08677304
; Patent No. 5721212
; GENERAL INFORMATION:
; APPLICANT: Wakefield, Thomas W.
; APPLICANT: Andrews, Philip C.
; APPLICANT: Stanley, James C.
; TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
; TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
; TITLE OF INVENTION: ANTICOAGULATION REVERSAL
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benita J. Rohm, Esq.
; STREET: 512 Springfield Avenue
; CITY: Cranford
; STATE: New Jersey
; COUNTRY: United States of America
; ZIP: 07016-1811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 6; ASCII (DOS)Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/677,304
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/152,488
; FILING DATE: 12-NOV-1993
; APPLICATION NUMBER: PCT/US92/08069
; FILING DATE: 14-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rohm, Benita J.
; REGISTRATION NUMBER: 28,664
; REFERENCE/DOCKET NUMBER: RM-7WG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-276-3344
; TELEFAX: 908-276-5543
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. 5721212 Relevant
; TOPOLOGY: No. 5721212 Relevant
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: N/A
; PUBLICATION INFORMATION:
; TITLE: N/A
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: PCT/US92/08069
; FILING DATE: 14-AUG-1993
; US-08-677-304-3

Query Match 2.0%; Score 7; DB 1; Length 29;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 AKAKAKA 255
DB 4 AKAKAKA 10

RESULT 10
US-08-436-703B-8
; Sequence 8, Application US/08436703B
; Patent No. 5919761
; GENERAL INFORMATION:
; APPLICANT: Wakefield, Thomas W.
; APPLICANT: Andrews, Philip C.
; APPLICANT: Stanley, James C.
; TITLE OF INVENTION: NOVEL PEPTIDES FOR
; TITLE OF INVENTION: HEPARIN AND LOW MOLECULAR
; TITLE OF INVENTION: WEIGHT HEPARIN
; TITLE OF INVENTION: ANTICOAGULATION REVERSAL
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benita J. Rohm, Esq.
; STREET: 6601 Woodward Avenue
; STREET: Suite 1525
; CITY: Detroit
; STATE: Michigan
; COUNTRY: United States of America
; ZIP: 48226
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk 1.44mb, 3.5"
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 6;
; SOFTWARE: ASCII (DOS)Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,703B
; FILING DATE: 08-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: N/A
; FILING DATE: N/A
; ATTORNEY/AGENT INFORMATION:
; NAME: Rohm, Benita J.
; REGISTRATION NUMBER: 28,664
; REFERENCE/DOCKET NUMBER: 7WK-060548-00233
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 313-965-1976
; TELEFAX: 313-965-1951
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: N/A
; PUBLICATION INFORMATION:
; TITLE: N/A
; PUBLICATION INFORMATION:
; US-08-436-703B-8

Query Match 2.0%; Score 7; DB 2; Length 29;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 AKAKAKA 255
DB 4 AKAKAKA 10

RESULT 11
US-08-995-172-18
; Sequence 18, Application US/08995172B
; Patent No. 6218112
; GENERAL INFORMATION:
; APPLICANT: Thatcher, David R

APPLICANT: Wilks, Paula E
TITLE OF INVENTION: Optimization of Gene Delivery and Gene Delivery Systems
FILE REFERENCE: CAC00026
CURRENT APPLICATION NUMBER: US/08/995.172B
CURRENT FILING DATE: 1997-12-22
EARLIER APPLICATION NUMBER: 60/033,908
EARLIER FILING DATE: 1996-12-23
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 18
LENGTH: 35
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: UNSURE
LOCATION: (35)
OTHER INFORMATION: xaa is cys with acm sidechain
FEATURE: INFORMATION: Description of Artificial Sequence: Peptide
US-08-995-172-18

Query Match 2.0%; Score 7; DB 4; Length 35;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 249 AKAKAKA 255
|||||||
DB 7 AKAKAKA 13

RESULT 12
US-08-993-008A-5
Sequence 5, Application US/08993008A
Patent No. 6153596
GENERAL INFORMATION:
APPLICANT: Liotta, Dennis C.
APPLICANT: Petros, John A.
APPLICANT: Wey, Shlow-Jyi
APPLICANT: Kair, Joan F.
APPLICANT: Pohl, Jan
TITLE OF INVENTION: Polycationic Oligomers
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: CO
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,008A
FILING DATE: 18-DEC-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/032,436
FILING DATE: 18-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sullivan, Sally A.
REGISTRATION NUMBER: 32,064
REFERENCE/DOCKET NUMBER: 33-95
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303-499-8080
TELEFAX: 303-499-8089
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid

STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-993-008A-5

Query Match 2.0%; Score 7; DB 4; Length 48;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 249 AKAKAKA 255
|||||||
DB 2 AKAKAKA 8

RESULT 13
US-08-993-008A-6
Sequence 6, Application US/08993008A
Patent No. 6153596

GENERAL INFORMATION:
APPLICANT: Liotta, Dennis C.
APPLICANT: Petros, John A.
APPLICANT: Wey, Shlow-Jyi
APPLICANT: Kair, Joan F.
APPLICANT: Pohl, Jan
TITLE OF INVENTION: Polycationic Oligomers
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: CO
COUNTRY: US
ZIP: 80303

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,008A
FILING DATE: 18-DEC-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/032,436
FILING DATE: 18-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sullivan, Sally A.
REGISTRATION NUMBER: 32,064
REFERENCE/DOCKET NUMBER: 33-95
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303-499-8080
TELEFAX: 303-499-8089
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 56 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-993-008A-6

Query Match 2.0%; Score 7; DB 4; Length 56;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 249 AKAKAKA 255
|||||||

Db 2 AKAKAKA 8

RESULT 14

US-08-769-211-2
Sequence 2, Application US/08769211

Patent No. 5830852

GENERAL INFORMATION:

APPLICANT: Thatcher et al.

TITLE OF INVENTION: Improved Pharmaceutical Compositions

TITLE OF INVENTION: Insulin-Receptor Mediated Gene Therapy.

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner & Witcoff, Inc.

STREET: 75 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1807

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Wordperfect 6.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/769,211

FILING DATE: 18-Dec-1996

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Kathleen M. Williams

REGISTRATION NUMBER: 34,380

REFERENCE/DOCKET NUMBER: 3255/02514

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-345-9100

TELEFAX: 617-345-9111

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 60 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-769-211-2

Query Match 2.0%; Score 7; DB 2; Length 60;

Best Local Similarity 100.0%; Pred. No. 19;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 AKAKAKA 255

DB 23 AKAKAKA 29

US-08-995-172-6

Sequence 6, Application US/08995172B

Patent No. 6218112

GENERAL INFORMATION:

APPLICANT: Thatcher, David R

APPLICANT: Wilks, Paula E

TITLE OF INVENTION: Optimization of Gene Delivery and Gene Delivery Systems

FILE REFERENCE: CACO0026

CURRENT APPLICATION NUMBER: US/08/995,172B

CURRENT FILING DATE: 1997-12-22

EARLIER APPLICATION NUMBER: 60/033,908

EARLIER FILING DATE: 1996-12-23

NUMBER OF SEQ ID NOS: 25

SOFTWARE: Patentln Ver. 2.1

SEQ ID NO 6

LENGTH: 61
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Peptide
US-08-995-172-6

Query Match 2.0%; Score 7; DB 4; Length 61;

Best Local Similarity 100.0%; Pred. No. 19;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 AKAKAKA 255

DB 24 AKAKAKA 30

Search completed: September 22, 2002, 18:32:15
Job time: 343 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 18:19:57 ; Search time 104.53 Seconds
(without alignments)
911.058 Million cell updates/sec

Title: US-09-898-216-1

Perfect score: 1767
Sequence: 1 MLRAARGHWGPFRAEGLSTG.....RDVQGTDSXDDELRYKMS 356

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 947138 seqs, 267508082 residues

Total number of hits satisfying chosen parameters: 947138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCIT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1763	99.8	356	US-09-898-216-1	Sequence 1, Appl
2	1646	93.2	356	US-09-629-469A-11010	Sequence 11010, A
3	1646	93.2	356	US-10-219-793-150	Sequence 150, App
4	1639	92.8	356	US-60-389-987-1904	Sequence 1904, Ap
5	1613.5	91.3	342	US-60-389-987-768	Sequence 768, App
6	1605	90.8	357	US-60-389-987-826	Sequence 826, App
7	1493	84.5	306	US-10-219-793-411	Sequence 411, App
8	619.5	35.1	275	US-10-219-999-39902	Sequence 39902, A
9	578.5	32.7	283	US-10-219-999-55341	Sequence 55341, A
10	573.5	32.5	263	US-10-219-999-51693	Sequence 51693, A
11	472	26.7	381	US-09-898-216-6	Sequence 6, Appl
12	466	26.4	305	US-09-791-537-98751	Sequence 98751, A
13	444.5	25.2	432	US-09-602-777A-140	Sequence 140, App
14	376.5	21.3	253	US-09-791-537-75409	Sequence 75409, A
15	376.5	21.3	576775	US-09-458-180-2	Sequence 2, Appl
16	354.5	20.1	199	US-09-791-537-98761	Sequence 98761, A
17	354.5	20.1	199	US-09-898-216-7	Sequence 7, Appl
18	292	16.7	60	US-10-219-793-418	Sequence 418, App
19	292	16.5	60	US-10-219-793-416	Sequence 416, App
20	281.5	15.9	288	US-09-898-216-3	Sequence 3, Appl
21	281.5	15.9	288	US-10-223-087-338	Sequence 338, App
22	281.5	15.9	288	US-10-223-087-338	Sequence 338, App
23	281.5	15.9	288	US-10-223-082-338	Sequence 338, App
24	281.5	15.9	288	US-10-223-083-338	Sequence 338, App
25	281.5	15.9	288	US-10-223-084-338	Sequence 338, App
26	281.5	15.9	288	US-10-223-085-338	Sequence 338, App

27	281.5	15.9	288	US-10-223-088-338	Sequence 338, App
28	281.5	15.9	288	US-10-223-090-338	Sequence 338, App
29	271	15.3	280	US-09-898-216-4	Sequence 4, Appl
30	263.5	14.9	347	PCIT-US02-13142-3455	Sequence 3455, Ap
31	263.5	14.9	347	US-10-128-714-3455	Sequence 3455, Ap
32	261	14.8	50	US-10-219-793-420	Sequence 420, App
33	256.5	14.5	347	PCIT-US02-13142-3455	Sequence 8455, Ap
34	256.5	14.5	347	US-10-128-714-8455	Sequence 8455, Ap
35	255.5	14.5	359	US-10-179-131-6020	Sequence 6020, Ap
36	231	13.1	70	US-10-219-793-415	Sequence 415, App
37	224	12.7	46	US-10-219-793-426	Sequence 426, App
38	221	12.5	45	US-10-219-793-424	Sequence 424, App
39	211	11.9	45	US-10-219-793-422	Sequence 422, App
40	209.5	11.9	383	US-60-389-987-853	Sequence 853, App
41	185	10.5	1693	US-10-155-881-30739	Sequence 30739, A
42	182.5	10.3	336	US-10-219-999-53682	Sequence 53682, A
43	172.5	9.8	329	US-10-219-999-57141	Sequence 57141, A
44	171.5	9.7	300	US-10-219-999-55568	Sequence 55568, A
45	170.5	9.6	284	US-10-219-999-57280	Sequence 57280, A

ALIGNMENTS

RESULT 1
US-09-898-216-1
; Sequence 1, Application US/09898216
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL HUMAN MEMBRANE PROTEIN
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/898, 216
; FILING DATE: 02-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/781,562
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0181 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 356 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: Consensus
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-898-216-1
Query Match 99.8%; Score 1763; DB 5; Length 356;

Best Local Similarity 100.0%; Pred. No. 6e-120;
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Query 1 MLARAARGHMGFPAGE--LSTGFMPRSGRASSGLPRNTVVLVFPQGEAMVVERMGFRHRI 60
Db 1 MLARAARGHMGFPAGE--LSTGFMPRSGRASSGLPRNTVVLVFPQGEAMVVERMGFRHRI 60
Query 61 PGLNLLIPVLDRIYVQSLKEIYINVEQSAVTLNDVTLQIDGVLYLRIMDPYKASYG 120
Db 61 PGLNLLIPVLDRIYVQSLKEIYINVEQSAVTLNDVTLQIDGVLYLRIMDPYKASYG 120
Query 121 DEYAVTQALQOTTRSELGKLSXDKVFRERESLNASTYDAINQADWCGRICLRYEIKI 180
Db 121 DEYAVTQALQOTTRSELGKLSXDKVFRERESLNASTYDAINQADWCGRICLRYEIKI 180
Query 181 HVPVRKESMOMQVEARRKRRATVLESEGTRESAINVAEGKKAQIILASEAKAEQINQ 240
Db 181 HVPVRKESMOMQVEARRKRRATVLESEGTRESAINVAEGKKAQIILASEAKAEQINQ 240
Query 241 AGEASAVLAKAKAKAEIRILAAALTOHNGDAAASLTVAEQYVSASFSLAKDSNTILL 300
Db 241 AGEASAVLAKAKAKAEIRILAAALTOHNGDAAASLTVAEQYVSASFSLAKDSNTILL 300
Query 301 NGDVTSMVAQMGVYGTALTKAPVPGTPTDLSGSSRDVQGTDSXDELDYKMS 356
Db 301 NGDVTSMVAQMGVYGTALTKAPVPGTPTDLSGSSRDVQGTDSXDELDYKMS 356

RESULT 2
US-09-629-469A-11010
; Sequence 11010, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; FILE REFERENCE: ISOGAI, TAKAO
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 11010
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-629-469A-11010

Query Match 93.2%; Score 1646; DB 5; Length 356;
Best Local Similarity 95.3%; Pred. NO. 1.9e-111;
Matches 341; Conservative 2; Mismatches 11; Indels 4; Gaps 2;

Query 1 MLARAARGHMGFPAGE--LSTGFMPRSGRASSGLPRNTVVLVFPQGEAMVVERMGFRHRI 58
Db 1 MLARAARGHMGFPAGE--LSTGFMPRSGRASSGLPRNTVVLVFPQGEAMVVERMGFRHRI 58
Query 59 LEPGLNLLIPVLDRIYVQSLKEIYINVEQSAVTLNDVTLQIDGVLYLRIMDPYKASYG 118
Db 59 LEPGLNLLIPVLDRIYVQSLKEIYINVEQSAVTLNDVTLQIDGVLYLRIMDPYKASYG 118
Query 119 VEDPEYAVTQALQOTTRSELGKLSXDKVFRERESLNASTYDAINQADWCGRICLRYEIKI 178
Db 119 VEDPEYAVTQALQOTTRSELGKLSXDKVFRERESLNASTYDAINQADWCGRICLRYEIKI 178
Query 179 DIHVPVRKESMOMQVEARRKRRATVLESEGTRESAINVAEGKKAQIILASEAKAEQIN 238
Db 179 DIHVPVRKESMOMQVEARRKRRATVLESEGTRESAINVAEGKKAQIILASEAKAEQIN 238
Query 239 QAAGEASAVLAKAKAKAEIRILAAALTOHNGDAAASLTVAEQYVSASFSLAKDSNTILL 298
Db 239 QAAGEASAVLAKAKAKAEIRILAAALTOHNGDAAASLTVAEQYVSASFSLAKDSNTILL 298
Query 299 PSNPGDVTSMVAQMGVYGTALTKAPVPGTPTDLSGSSRDVQGTDSXDELDYKMS 356
Db 299 PSNPGDVTSMVAQMGVYGTALTKAPVPGTPTDLSGSSRDVQGTDSXDELDYKMS 356

RESULT 3
US-10-219-793-150
; Sequence 150, Application US/10219793
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; FILE REFERENCE: P2008PICI
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 09/209,462
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: PCT/US98/12125
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/049,547
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,548
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,549
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,550
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,566
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,606
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,607
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,608
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,609
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,610
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,611
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/050,901
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/052,989
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058,665
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,668
; PRIOR FILING DATE: 1997-09-12

```

; FILE REFERENCE: 660088.465P2
; CURRENT APPLICATION NUMBER: US/60/389,987
; CURRENT FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 3025
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1904
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-389-987-1904

Query Match          92.8% DB 7; Length 356;
Best Local Similarity 95.0% Pred. No. 6e-111;
Matches 340; Conservative 2; Mismatches 12; Indels 4; Gaps 2.

QY      1 MLARAARCHMGPFAEG--LSTGFMFRSGRASSGGLPRTNVIVLVPPOOEAWVERMGRFHRI 58
        ||||| 1 : : : || ||||| ||||| ||||| ||||| |||||
Db      1 MLARAARTGALLRGSLASGRAP--RASSGGLPRTNVIVLVPPOEAWVERMGRFHHI 58

QY      59 LEPGLNIIIPVLDRIRYVQSLKEIYINVPESAVTLDNVTLOIDGVLYLRIMDPKASYG 118
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      59 LEPGLNIIIPVLDRIRYVQSLKEIYINVPESAVTLDNVTLOIDGVLYLRIMDPKASYG 118

QY      119 VEDPEYANTOLAQOTMRSSELGKLSDKYFRERRESINASTIVAINAQAACWGRCLRYETK 178
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      119 VEDPEYANTOLAQOTMRSSELGKLSDKYFRERRESINASTIVAINAQAACWGRCLRYETK 178

QY      179 DIHPPRVKESMOMQVEARRRRRAVTLSEGTRESAINVADEKKQAQILASEAKAEQIN 238
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      179 DIHPPRVKESMOMQVEARRRRRAVTLSEGTRESAINVADEKKQAQILASEAKAEQIN 238

```

[illegible]

```

Oy 77 QSLKEIVINPEQSAVTLDNVTLQIDGVLYLRIMDPYKASGVEDBEYAVTQLOATTMRS 136
Db 63 QSLKEIVINPEQSAVTLDNVTLQIDGVLYLRIMDPYKASGVEDBEYAVTQLOATTMRS 122
Oy 137 ELGLSKAKVYRERESLNASIVDAINQADCMGIRCLREYKIDHPPRVESMOMOEVA 196
Db 123 ELGLSKLDKVRERESLNASIVDAINQADCMGIRCLREYKIDHPPRVESMOMOEVA 182
Oy 197 ERRKRATVLESEGTRESAIVNAEGKKQAOILASEAKAQINQAGSAVLAQAKAKAE 256
Db 183 ERRKRATVLESEGTRESAIVNAEGKKQAOILASEAKAQINQAGSAVLAQAKAKAE 242
Oy 257 AIRILAAALTOHNGDAASLTVAEOYVSAFSLKADSNITLLPSNGDYTSVNAQAMGY 316
Db 243 AIRILAAALTOHNGDAASLTVAEOYVSAFSLKADSNITLLPSNGDYTSVNAQAMGY 302
Oy 317 GALTAPVPGTPDLSGSSRDVQGTDAASXDELDVRKMS 356
Db 303 GALTAPVPGTPDLSGSSRDVQGTDAASXDELDVRKMS 342

RESULT 6
US-60-389-987-826
Sequence 826, Application US/60389987
GENERAL INFORMATION:
APPLICANT: Ghosh, Soumitra S.
APPLICANT: Fahy, Eoin D.
APPLICANT: Zhang, Bing
APPLICANT: Gibson, Bradford W.
APPLICANT: Taylor, Steven W.
APPLICANT: Glenn, Gary M.
APPLICANT: Warnock, Dale E.
TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
FILE REFERENCE: 660088, 465P2
CURRENT APPLICATION NUMBER: US/60/389,987
CURRENT FILING DATE: 2002-06-17
NUMBER OF SEQ ID NOS: 3025
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 826
LENGTH: 357
TYPE: PRT
ORGANISM: Homo sapiens
US-60-389-987-826

Query Match          90.8%: Score 1605; DB 7; Length 357;
Best Local Similarity 93.8%: Pred No. 1,8e-108;
Matches 332; Conservative 2; Mismatches 4; Indels 16; Gaps 2;

Oy 17 LSTGFWPRSGRASSGLPRTTVLVLPQOEAMVVERMGFRHRLLEPILILIPVLDRIYV 76
Db 6 LASGRAPR--RASSGLPRTTVLVLPQOEAMVVERMGFRHRLLEPILILIPVLDRIYV 63
Oy 77 QSLKEIVINPEQSAVTLDNVTLQIDGVLYLRIMDPYKASGVEDBEYAVTQLOATTMRS 136
Db 64 QSLKEIVINPEQSAVTLDNVTLQIDGVLYLRIMDPYKASGVEDBEYAVTQLOATTMRS 123
Oy 137 ELGLSKXDKVFERESLNASIVDAINQADCMGIRCLREYKIDHPPRVESMOMQ--- 193
Db 124 ELGLSKLDKVRERESLNASIVDAINQADCMGIRCLREYKIDHPPRVESMOMQVGA 183
Oy 194 -----VEARRKATVLESEGTRESAIVNAEGKKQAOILASEAKAQINQAG 242
Db 184 KECMEKGLRAPVEARRKATVLESEGTRESAIVNAEGKKQAOILASEAKAQINQAG 243
Oy 243 EASAVLAKAKAKAEAIRILAAALTOHNGDAASLTVAEOYVSAFSLKADSNITLLPSNP 302
Db 244 EASAVLAKAKAKAEAIRILAAALTOHNGDAASLTVAEOYVSAFSLKADSNITLLPSNP 303
Oy 303 GDYTSVNAQAMGYGALTAPVPGTPDLSGSSRDVQGTDAASXDELDVRKMS 356
Db 303 GDYTSVNAQAMGYGALTAPVPGTPDLSGSSRDVQGTDAASXDELDVRKMS 356
```

```

Db 304 GDYTSVNAQAMGYGALTAPVPGTPDLSGSSRDVQGTDAASXDELDVRKMS 357

RESULT 7
US-10-219-793-411
Sequence 411, Application US/10219793
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 86 Human Secreted Proteins
FILE REFERENCE: P2008P1C1
CURRENT APPLICATION NUMBER: US/10/219,793
CURRENT FILING DATE: 2002-08-16
PRIOR APPLICATION NUMBER: 09/209,462
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: PCT/US98/12125
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/049,547
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,548
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,549
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,550
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,566
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,606
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,607
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,608
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,609
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,610
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/049,611
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/050,901
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/052,989
PRIOR FILING DATE: 1997-06-13
PRIOR APPLICATION NUMBER: 60/051,919
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/055,984
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/058,665
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,668
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,669
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,750
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,971
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,972
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,975
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/060,834
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,841
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,844
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/060,865
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/061,059
PRIOR FILING DATE: 1997-10-02
PRIOR APPLICATION NUMBER: 60/061,060
PRIOR FILING DATE: 1997-10-02
NUMBER OF SEQ ID NOS: 737
```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 411
; LENGTH: 306
; TYPE: prt
; ORGANISM: Homo sapiens
US-10-219-793-411

```

```

Query Match      84.5%: Score 1493; DB 6; Length 306;
Best Local Similarity 99.3%: Pred. No. 1.9e-100;
Matches 304; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 51 RMGRFRRIIEPGINILIPVLDRIYVQSLKEIYINPEQSAVTLNDVTLQIDGVLYLRIM 110
DB 1 RMGRFRRIIEPGINILIPVLDRIYVQSLKEIYINPEQSAVTLNDVTLQIDGVLYLRIM 60
QY 111 DPKASYGVEDEPYAVTOLAQTMRSELGKLSDKYFRERESINASIVDAINQADCGWI 170
DB 61 DPKASYGVEDEPYAVTOLAQTMRSELGKLSDKYFRERESINASIVDAINQADCGWI 120
QY 171 RCLRYEIKIHVPVRKESMOMOVEARRRRATVLESEGTRESAINVAEGKKQAQIILASE 230
DB 121 RCLRYEIKIHVPVRKESMOMOVEARRRRATVLESEGTRESAINVAEGKKQAQIILASE 180
QY 231 AERAEQINQAGEASAVLAKAKAKAEIRILAAALTOHNGDAASLTVAEQYSAFSKLA 290
DB 181 AERAEQINQAGEASAVLAKAKAKAEIRILAAALTOHNGDAASLTVAEQYSAFSKLA 240
QY 291 KQSNITLFSNPGDVTSMVAQAMGVGALTAPVPGTPOSLSGSSRDVQGTDAKXDEEL 350
DB 241 KQSNITLFSNPGDVTSMVAQAMGVGALTAPVPGTPOSLSGSSRDVQGTDAKXDEEL 300
QY 351 DRYKMS 356
DB 301 DRYKMS 306

```

```

RESULT 8
US-10-219-999-39902

```

```

; Sequence 39902, Application US/10219999
; GENERAL INFORMATION:

```

```

; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)C
; CURRENT APPLICATION NUMBER: US/10/219,999
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/324,109
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/312,544
; PRIOR FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 63520
; SEQ ID NO 39902
; LENGTH: 275
; TYPE: prt
; ORGANISM: Zea mays
US-10-219-999-39902

```

```

Query Match      35.1%: Score 619.5; DB 6; Length 275;
Best Local Similarity 48.2%: Pred. No. 4.1e-37;
Matches 131; Conservative 47; Mismatches 77; Indels 17; Gaps 2;

```

```

QY 100 QIDGVLYLRIMDPYKASYGVEDEPYAVTOLAQTMRSELGKLSXKDYFRERESINASIYD 159
DB 1 QIDSVIYKIMDYLRASYGVENIYAVLQAOQTMRSSELGKITLDTFEERDALNEKIVS 60
QY 160 AINQADCGWICIRLREIKIHVPVRKESMOMOVEARRRRATVLESEGTRESAINVAE 219

```

```

DB 61 AINEATDMGICIRYEIRIDINPPACIRQAMENQAEERKKRAQIILESEGMQAOIILESE 120
QY 220 GKQAOIILASEAKAQINQAGEASAVLAKAKAKAEIRILAAALTOHNGDAASLTVA 279
DB 121 GKQAOIILASEGAMLDLANAKGAABEIRILAAALTOHNGDAASLTVAEQYSAFSKLA 180
QY 280 EQYVSAFSKLAQDSNTILPNSNPGDVTSMVAQAMGVGALTAPVPGTPOSLSGSSRDVQGTDAKXDEELDRV 354
DB 181 EQYVSAFSKLAQDSNTILPNSNPGDVTSMVAQAMGVGALTAPVPGTPOSLSGSSRDVQGTDAKXDEELDRV 240
QY 324 -VPGTPOSLSGSSRDVQGTDAKXDEELDRV 354
DB 241 ETPAPSSSEASKTPPLEEADSNQTFSLQRPK 272

```

```

RESULT 9
US-10-219-999-55341

```

```

; Sequence 55341, Application US/10219999
; GENERAL INFORMATION:

```

```

; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)C
; CURRENT APPLICATION NUMBER: US/10/219,999
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/324,109
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/312,544
; PRIOR FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 63520
; SEQ ID NO 55341
; LENGTH: 283
; TYPE: prt
; ORGANISM: Zea mays
US-10-219-999-55341

```

```

Query Match      32.7%: Score 578.5; DB 6; Length 283;
Best Local Similarity 53.4%: Pred. No. 4e-34;
Matches 119; Conservative 37; Mismatches 50; Indels 17; Gaps 2;

```

```

QY 14 AEGLSGTFMWRSGRASSGL-----PRNTYVLYFPQOEAHVVERMGFRHILREGLNILI 67
DB 72 APALSRFSFNFNRDSSMDEPPPEVPMGVSIPEKKAVVERFGKYLTLGSGFHLI 131
QY 68 PVLDRIRYVQSLKEIYINPEQSAVTLNDVTLQIDGVLYLRIMDPYKASYGVEDEPYAVT 127
DB 132 PAVDRIRYVQSLKEIYINPEQSAVTLNDVTLQIDGVLYLRIMDPYKASYGVEDEPYAVT 191
QY 128 QLAQTMRSSELGKLSDKYFRERESINASIVDAINQADCGWICIRLREIKIHVPVRK 187
DB 192 QLAQTMRSSELGKLTLDTFEERDALNEKIVSAINEATDMGICIRYGIIRDINPPACIR 251
QY 188 ESMQOMVEARRRRATVLESEGTRESAINVAEGKKQAQIILASE 230
DB 252 QAMEQOAEARRRRATVLESEGTRESAINVAEGKKQAQIILASE 283

```

```

RESULT 10
US-10-219-999-51693

```

```

; Sequence 51693, Application US/10219999
; GENERAL INFORMATION:

```

```

; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 17:26:03 ; Search time 44.59 Seconds

(without alignments)
767.163 Million cell updates/sec

Title: US-09-898-216-1

Perfect score: 1767

Sequence: 1 MLRAARGHWGPFRAEGLSTG.....RDVQGTDSXDELDPRVYKMS 356

Scoring table: BLOSUM62

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database : PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
1	1605	90.8	357	2 T02246	hypothetical prote
2	957.5	54.2	334	2 T21562	hypothetical prote
3	861.5	46.8	354	2 T38599	conserved hypotet
4	826	46.7	515	2 T05863	hypothetical prote
5	811.5	45.9	312	2 D97756	hypothetical prote
6	806	45.6	311	2 B71689	hypothetical prote
7	722	40.9	374	2 T18493	hypothetical prote
8	716	40.5	315	2 H81907	probable periplasm
9	712.5	40.3	315	2 F81107	stomatol/Mec-2 fam
10	499.5	28.3	321	2 S74505	erythrocyte band 7
11	497	28.1	344	2 D75311	conserved hypotet
12	496.5	28.1	337	2 A82838	conserved hypotet
13	486	27.5	310	2 E87667	band 7/Mec-2 fam1
14	482.5	27.3	322	2 A12114	hypothetical prote
15	477	27.0	304	2 AB0375	conserved hypotet
16	475	26.9	319	2 T36857	conserved hypotet
17	472	26.7	381	2 C70711	hypothetical prote
18	468	26.5	305	2 AH0564	probable membrane
19	467.5	26.5	306	2 F82556	conserved hypotet
20	466	26.4	305	2 H64779	probable membrane
21	466	26.4	305	2 H90697	probable proteins
22	466	26.4	305	2 B85548	probable proteins
23	465	26.3	350	2 A98264	hypothetical 33.7K
24	465	26.3	350	2 AH3020	conserved hypotet
25	460	26.0	298	2 A71159	probable membrane
26	460	26.0	299	2 F75003	stomatol-like prot
27	457	25.9	318	2 C69192	stomatol-like prot
28	454.5	25.7	328	2 A13511	stomatol like prot
29	444.5	25.2	374	2 C87134	conserved hypotet

30	442.5	25.0	322	2 A97030	membrane protease
31	423.5	24.0	305	2 E72322	conserved hypotet
32	417.5	23.6	392	2 G84191	hypothetical prote
33	412	23.3	309	2 D90520	conserved hypotet
34	376.5	21.3	253	2 F70378	erythrocyte band 7
35	369	20.9	278	2 AC2354	hypothetical prote
36	362	20.5	249	2 C69427	membrane protein h
37	354.5	20.1	199	2 C64403	membrane protein h
38	345.5	19.6	278	2 D72522	probable erythrocy
39	339	19.2	297	2 G69104	stomatol-like prot
40	327.5	18.5	266	2 C71027	probable erythrocy
41	324.5	18.4	264	2 A83590	probable stomatin-
42	324	18.3	256	2 B95981	probable stomatin-
43	323.5	18.3	262	2 T34691	hypothetical prote
44	321.5	18.2	268	2 B75107	stomatol-like prot
45	310	17.5	267	2 H90388	erythrocyte band 7

ALIGNMENTS

RESULT 1
T02246
hypothetical protein P1.11659_4 - human
C:Species: Homo sapiens (man)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 11-Jan-2000
C:Accession: T02246
R:lamerdin, J.E.; McCreedy, P.M.; Skowronski, E.; Adamson, A.W.; Burthart-Schultz, K.; Christensen, M.; Georgescu, A.; Avila, J.; Liu, S.; Atlix, C.; Andreise, F.; Trank
submitted to the EMBL Data Library, March 1998
A:Authors: Duarte, S.; Lucas, S.; Bruce, R.; Thomas, P.; Quan, G.; Krommiller, B.; Ar
A:Description: Sequence analysis of a human P1 clone containing the XRC9 DNA repair
A:Reference number: 214637
A:Accession: T02246
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-357 <LAN>
A:Cross-references: EMBL:AC004472; NID:g2984582; PIDN:AAC07983.1; PID:g2984585
C:Genetics:
A:Map position: 9
A:introns: 2/3; 48/3; 82/1; 101/3; 135/3; 194/3; 243/1; 269/3; 312/3
A:Superfamily: erythrocyte band 7 integral membrane protein

Query Match	Score	Length	DB	2:	Indels	Gaps
Best Local Similarity	93.8%	Pred. No. 2.9e-98;				
Matches	332;	Conservative	2;	Mismatches	4;	Indels 16; Gaps 2;
QY	17	LSTGFWRSGRASGLPRTVTLFVPOQAWYERMGFRHILEPGLNLIIPVLRIRYV	76			
DB	6	LASGRAPR--RASSGLPRTVTLFVPOQAWYERMGFRHILEPGLNLIIPVLRIRYV	63			
QY	77	OSLKEIVNPEQSAVTLQIDGVLYLRIMDPYKASYGVDEPVAVTOLAOTMRS	136			
DB	64	OSLKEIVNPEQSAVTLQIDGVLYLRIMDPYKASYGVDEPVAVTOLAOTMRS	123			
QY	137	ELGKLSXDFVPRRESLNASTVDATINQADCGIRCLREIKOIHVPYKESMOMQ---	193			
DB	124	ELGKLSXDFVPRRESLNASTVDATINQADCGIRCLREIKOIHVPYKESMOMQVCA	183			
QY	194	-----YEAERRKRAVYLESEGTRESAINVAEGKKQAOIILASEAKAEQIQNAG	242			
DB	184	KEGWKGLAPYEAERKRAVYLESEGTRESAINVAEGKKQAOIILASEAKAEQIQNAG	243			
QY	243	ESAVYLAKAKAKAEATRIITAAALTOHNGDAASLYTAEQYVSASFSLADSNITLLPSNP	302			
DB	244	ESAVYLAKAKAKAEATRIITAAALTOHNGDAASLYTAEQYVSASFSLADSNITLLPSNP	303			
QY	303	GDVTSVVAQAMGVYALTKAPVPGTDSLSGSSRRVVGCTDASXDELDPRVYKMS	356			
DB	304	GDVTSVVAQAMGVYALTKAPVPGTDSLSGSSRRVVGCTDASXDELDPRVYKMS	357			

RESULT 2
 T21562
 hypothetical protein F30A10.5 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
 C:Accession: T21562
 R:Barlow, K.
 submitted to the EMBL Data Library, October 1996
 A:Reference number: Z19442
 A:Accession: T21562
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-334 <MIL>
 A:Cross-references: EMBL:Z81072; PIDN:CAB03018.1; GSPDB:GN00019; CESP:F30A10.5
 A:Experimental source: clone F30A10
 C:Genetics:
 A:Gene: CESP:F30A10.5
 A:Map position: 1
 A:Introns: 16/2; 68/3; 124/3; 227/2
 C:Superfamily: erythrocyte band 7 integral membrane protein

Query Match	54.28;	Score 957.5;	DB 2;	Length 334;
Best Local Similarity	61.28;	Pred. No. 9.6e-56;		
Matches 188;	Conservative 58;	Mismatches 46;	Indels 15;	Gaps 2

[illegible]

RESULT 3
T39599
conserved hypothetical protein SPBC16G5.07c - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Feb-2000
C:Accession: T39599
R:Llyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.
submitted to the EMBL Data Library, May 1998
A:Reference number: Z21866
A:Accession: T39599
A:Status: Preliminary: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-354 <LYN>
A:Cross-references: EMBL:AL023554; PTDN:CA119027.1; GSPDB:GN00067; SPDB:SPBC16G5.07c
A:Experimental source: strain 972h-, cosmid c16G5
C:Genetics:
A:Gene: SPDB:SPBC16G5.07c
A:Map position: 2
A:Introns: 72/2; 265/2
C:Superfamily: erythrocyte band 7 integral membrane protein

Query Match	48.8%;	Score 861.5;	DB 2;	Length 354;
Best Local Similarity	56.7%;	Pred. No. 2.1e-49;		
Matches 170;	Conservative 56;	Mismatches 71;	Indels 3;	Gaps 2;

[illegible]

RESULT 4

T05863

hypothetical protein T29A15.70 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 13-Aug-1999

C:Accession: T05863

R:Bevan, M.; Van Der Schuren, J.; Chang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.;

submitted to the Protein Sequence Database, March 1999

A:Reference number: Z15455

A:Accession: T05863

A:Molecule type: DNA

A:Residues: 1-515 <BEV>

A:Cross-references: EMBL:AL035602

C:Experimental source: cultivar Columbia; BAC clone T29A15

C:Genetics:

A:Map position: 4

A:Introns: 52/2; 133/3; 183/3; 203/1; 234/1; 267/1; 297/3; 317/3; 394/1; 451/3

A:Note: T29A15.70

[illegible]

A:Accession: D75311
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-344 <NH2>
A:Cross-references: GB:AE002048; GB:AE000513; NID:96459929; PIDN:AAF11687.1; PID:9645993
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2143
A:Map position: 1
C:Superfamily: erythrocyte band 7 integral membrane protein

Query Match 28.1%; Score 497; DB 2; Length 344;
Best Local Similarity 37.2%; Pred. No. 1.9e-25;
Matches 108; Conservative 67; Mismatches 103; Indels 12; Gaps 5;
QY 41 VQGEAWYVERNGRFRHILEPGLNIIIPVLDRIYVOSLKEIYINPEQSAVTLDNVTLQ 100
DB 25 VQGEAWYVERNGRFRHILEPGLNIIIPVLDRIYVOSLKEIYINPEQSAVTLDNVTLQ 84
QY 101 IDGVYLRIIMDPYKASVGEDEPEYAVTOLAOTMRSELGKLSXDKYFRERESINASTIVA 160
DB 85 VDAVYFYOVLDAKASVYERNLEQAVLNLITNIRTVTSMDELINRKTINAKLVV 144
QY 161 INQADCMGIRCLRYEIKDINHPPRYKESMOMQVEAERKRRAVLESECTFESAINVAG 220
DB 145 VDEAREPMGVKTVTRIEVKDKIKPPADLVASMAROMKAREKRANILDAEGFQAALIKADG 204
QY 221 KQAOILASAEKAEQINOAA-GEASAVLAKAKAKAEAIRIILAAITQHNQDAASLTVA 279
DB 205 EKQAVALKAEKGRK----QASFMSEARERRRAQAEAEATRVVSOALFAGCNVQ-AYNYFLA 258
QY 280 EGYVSFSLAKDSN--TILPSNPGDVTSMVAQAMGVGALTKAPVPGT 327
DB 259 QQYVEALRPVVASAPNOKTILPI---EATSIIGSLGIGIAEVAKKATPES 305

RESULT 12
A82838
conserved hypothetical protein XF0185 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C:Accession: A82838
C:Anonymous: The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequences
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: A82838
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-337 <SIM>
A:Cross-references: GB:AE003872; GB:AE003849; NID:g9104975; PIDN:AAF2998.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Stimpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
B:Jones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carreir, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, B
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmeri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsuhako, M.H.; Vallada, H.; Van Silvas, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0185
C:Superfamily: erythrocyte band 7 integral membrane protein

Query Match 28.1%; Score 496.5; DB 2; Length 337;
Best Local Similarity 36.2%; Pred. No. 1.9e-25;
Matches 119; Conservative 67; Mismatches 116; Indels 27; Gaps 8;
QY 30 SGLPRNTV-----VLFVPOGEAWYVERNGRFRHILEPGLNIIIPVLDRIY 74
DB 19 SMLPNVLAIVLVAGVILLESKSVIMVPOGEAWYVERNGRFRHILEPGLNIIIPVLDRIY 78
QY 75 YVOSLKEIYINPEQSAVTLDNVTLQIDGVYLRIIMDPYKASVGEDEPEYAVTOLAOTMR 134
DB 79 RYVSMEOVLAVPSQEVITIKDNAGVGVDEYVFFQVILDAKAAEAVNLITMALVOTNI 138
QY 135 RSELGKLSXDKYFRERESINASTIVAIDINOACMGIRCLRYEIKDINHPPRYKESMOMQV 194
DB 139 RYVGSIDPEDESQRETIINAKLISYEHAISPMGVKTVTRIEVKDKIKPPADLVASMAROMKAREKRANILDAEGFQAALIKADG 198
QY 195 EERRRRATVLESECTFESAINVAGKQAOILASAEKAEQINOAA-GEASAVLAKAKAKAEAIRIILAAITQHNQDAASLTVA 254
DB 199 MAEQTRATILLEGVROSAIIRLADGEKQAAVYEAEGREAEAFRDA--EARELAEAEAK 256
QY 255 AEAIRILAAITQHNQDAASLTVAEOYVSAPSKLAKDSN--TILPSNPGDVTSMVAQAMGVGALTKAPVPGT 312
DB 257 --ATRLISEAIIISQGNVQ-AVNYFVAKYVEAEKELATAPNOKFILMPETSGIISIA-- 311
QY 313 MGYYGALTKAPVPGTSPDLSGSSRPDVG 341
DB 312 -GI-ADLTREAL-GKSDHATPSQREHRG 337

RESULT 13
E87667
band 7/Mec-2 family protein [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: E87667
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: E87667
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-310 <STO>
A:Cross-references: GB:AE005673; NID:g13425081; PIDN:AAK25337.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC3375
C:Superfamily: erythrocyte band 7 integral membrane protein

Query Match 27.5%; Score 486; DB 2; Length 310;
Best Local Similarity 35.5%; Pred. No. 8.5e-25;
Matches 107; Conservative 79; Mismatches 97; Indels 18; Gaps 7;
QY 36 TVVLEFVPOGEAWYVERNGRFRHILEPGLNIIIPVLDRIYVOSLKEIYINPEQSAVTL 95
DB 5 SAKIIVPQGRREFVBERFGYTRILKIGITILPIPLETVGRVNMQGVLDVPOQEVITKD 64
QY 96 NVTLOIDGVYLRIIMDPYKASVGEDEPEYAVTOLAOTMRSELGKLSXDKYFRERESLNA 155
DB 65 NNSVKDAIVLFIOVMDAAAVFVDIMMYAIQLOAOTNIRTVVGADELVSQRAINS 124
QY 156 STYDANQADCMGIRCLRYEIKDINHPPRYKESMOMQVEAERKRRAVLESECTFESAIN 215
DB 125 RLSTIDHTGPGGVAVVAREIKDILTPPADITNAMRQMKAEERRAVITEAGGESQAOI 184
QY 216 NVAEGKQAOILASAEKAEQINOAGASAVLAKAKAKAEAIRIILAAITQHNQDA-AA 274
DB 185 AAREGQKQSAIILQAEGR-----EAAFRDAEAREAEAEKAKATAFVSEAIK--GQVNAI 238
QY 275 SLTVAEOYVSAPSKLAK--DSNTILPSNPGDVTSMVAQAMGVGALTK-----ADVPCT 327

Oy	156	SIVDAINAOADOCWGIRCLRYEIKLHVPRKVESOMQVEARRRRARVLEEGRRESAI	215
Dy	138	LILRLRELDISTDPMGKAVTVRELTRDIMPSSKAVIDSMELQWTAERKKRAALITSEGORDSFI	197
Oy	216	NVAEGRKQAOLIASAEKAEKAEQINOAGAENSAVLAAKAKAEBRILAALTQHN-GDAAA	274
Dy	198	NSAGDADAARVLEAEKAKKAILNMADEAQOKKVLEAKKATPAEALSTILETKLSDNHAREAL	257
Oy	275	SILTVAEOYVSASFSLAKDSNTLLPSNQGDTYSMAQAONGVG	317
Dy	258	QFLLAOQYLNMGTITGSSDSSKVMFLDPNLTISTLEGVRSIVG	300
<hr/>			
RESULT 2			
ID	ye88_myctu	STANDARD:	PRT: 381 AA.
AC	p71768;		
DT	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Hypothetical 41.2 kDa protein RV1488.		
CN	RV1488 OR MT1533.2 OR MTCY277.09.		
OS	Mycobacterium tuberculosis.		
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;		
CC	Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.		
NCBI_TaxID=1773;	[1]		
RN	SEQUENCE FROM N.A.		
RP	STRAIN=H37RV;		
RC	MEDLINE=98295987; PubMed=9634230;		
RX	Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,		
RA	Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,		
RA	Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,		
RA	Davies R., Jagels K., Felwell T., Gentles S., Hamlin N., Holroyd S.,		
RA	Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,		
RA	Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,		
RA	Rutler S., Seeger K., Skelton S., Squares S., Squares R.,		
RA	Sultun J.E., Taylor K., Whitehead S., Barrell B.G.;		
RT	"Deciphering the biology of Mycobacterium tuberculosis from the		
RL	complete genome sequence."		
LN	Nature 393:537-544(1998).		
RN	SEQUENCE FROM N.A.		
RP	STRAIN=CDC 1551 / Oshkosh;		
RC	Fleischmann R.D., Allard D., Eisen J.A., Carpenter L., White O.,		
RA	Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,		
RA	Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,		
RA	Delcher A., Utterback T., Weisman H., Kouri H., Gill J., Mikula A.,		
RA	Bisbal W.;		
RT	"Whole genome comparison of mycobacterium tuberculosis clinical and		
RL	laboratory strains"		
CC	Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.		
CC	-I- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@lsb-sib.ch).		
CC	-----		
DR	EMBL: Z79701; CAB02038.1; -		
DR	EMBL: AE007032; AAA45800.1; -		
DR	TIGR: MT1533.2; -		
DR	TubercuList; RV1488; -		
DR	InterPro: IPR001107; Band_7.		
DR	InterPro: IPR001972; Stomatin.		
DR	Pfam: PF01145; Band_7; 1.		
DR	SMART: SM00244; PHB-1.		
DR	PROSITE: PS01270; BAND_7; 1.		
DM	Hypothetical protein; Transmembrane; Complete proteome.		

FT	TRANSMEM	3	23	POTENTIAL
SQ	SEQUENCE	381 AA:	41282 MW:	32747A51D653CF91 CRC64:
	Query Match	26.7%:	Score 472:	DB 1: Length 381:
	Best Local Similarity	33.2%:	Pred. No. 1,2e-23:	
	Matches 110:	Conservative 72:	Mismatches 123:	Indels 26: Gaps
QY	38	VLEVPQEEAWVVERNGFRHLLPEQLNLLPVLDRIRYVGSKEIYINVPQSAVTLDNV	97	
DB	26	VALIIPQAEAAVIERLGRYSRTVSQGLTLLPFIIDRVRRARVLDREIRVVSFPQPIYTEDNL	85	
QY	98	TLQIDGVLYLRLMDPKYKSCVGEDEPEYAVNPQLQATMNSGEIIXSKVFFERESLNASI	157	
DB	86	TLNIDTIVYFQVYVQAAVYVEISNIVYGEDELTTTLTNVVGMTLQGLTTSRQINAOQL	143	
QY	158	VDAINQAADCGCIRCLARYEIKDINHPPVKKSSOMOVQEAERKRATVLESEGTRESAINV	217	
DB	146	RGVLEDEALGRNGRLRARARVELNSIDPPSIQASMCKMADEKRAMILLTAEGTPEAAIKQ	205	
QY	218	AEGKKQAO-----ILASEAKAEQINQAGASAVLAKAKAKAEAIRILAALT	266	
DB	206	AEGQKQAOIILAEAGAKQAAIILAEADRSRMILRAQGEFAAAYLOAQOGAAAIETKFAAI	264	
QY	267	QHNGDAASLTVAEQYVAFSEKLAK-DSNTI-LTPSNGDVTSNVAQAMGVG-----AL	319	
DB	265	--KAGRPPEMLAIQYLLQTLPEMARGDKNAKVVSPFNALOGFTRLGLKPGEDGVFRF	3222	
QY	320	TKAPVPGTPTDLSLSSGSSSRDVG-----TDAS	345	
DB	323	EPSPYEDQPKHAADDDAEVAGWFTSDTDP	353	
RESULT	3			
YBKB_ECOLI	STANDARD:	PRT:	305 AA.	
AC	P7367:			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Hypothetical protein ybkb.			
GN	YBKB OR B0489 OR Z0642 OR EGS0552.			
OS	Escherichia coli, and			
OS	Escherichia coli O157:H7.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Escherichia.			
OX	NCBI_TaxID=562, 83334;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-K12 / MG1655.			
RX	MEDLINE=97426617; PubMed=9278503.			
RA	Blattner F.R., Plunkett G. II, Bloch C.A., Perna N.T., Burland V.,			
RA	Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,			
RA	Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,			
RA	Mau B., Shao Y.;			
RT	"The complete genome sequence of Escherichia coli K-12.";			
RT	Science 277:1453-1474(1997).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RP	Roberts D., Allen E., Araujo R., Aparicio A., Chung E., Davis K.,			
RA	Duncan M., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,			
RA	Lew H., Lin D., Namath A., Oefner P., Schramm S., Davis R.W.;			
RL	Submitted (Jan-1997) to the EMBL/GenBank/DBJ databases.			
RC	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=O157:H7 / EDL933 / ATCC 700927;			
RX	MEDLINE=21074935; PubMed=11206551;			
RA	Perna N.T., Plunkett G. II, Burland V., Mau B., Glasner J.D.,			
RA	Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,			
RA	Posfal G., Heckett J., Klink S., Boutin A., Shao Y., Miller L.,			
RA	Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,			
RA	Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,			
RA	Weich R.A., Blattner F.R.;			

RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7." ;
 RL Nature 409:529-533(2001).
 [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; Pubmed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Robe T.,
 Iida T., Takami H., Honda T., Sasekawa C., Ogasawara N., Yasunaga T.,
 Kubura S., Shiba T., Hattori M., Shitagawa H.,
 RA "Complete genome sequence of enterohaemorrhagic *Escherichia coli*
 RT O157:H7 and genomic comparison with a laboratory strain K-12." ;
 RL DNA Res. 8:11-22(2001).
 CC -1- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC
 CC EMBL; AE000155; AAC73591.1; -
 DR EMBL; U82664; AAB40243.1; -
 DR EMBL; AE005230; AAC54846.1; -
 DR EMBL; AP002552; BAB33975.1; -
 DR Ecogene; EG13258; ypbk.
 DR InterPro; IPR001107; Band_7.
 DR InterPro; IPR001972; Stomatin.
 DR Pfam; PF01145; Band_7; 1.
 DR SMART; SM00244; PHB; 1.
 DR PROSITE; PS01270; BAND_7; 1.
 DR Hypothetical protein; Transmembrane; Complete proteome.
 KW TRANSMEM 3 23 POTENTIAL.
 FT SEQUENCE 305 AA: 33743 MW; 0A780F65790ACC3 CRC64;
 SQ
 Query Match 26.4%; Score 466; DB 1; Length 305;
 Best Local Similarity 34.8%; Pred. No. 2.2e-23;
 Matches 98; Conservative 76; Mismatches 94; Indels 14; Gaps 6;
 QY 38 VLEFPOEAWVERMGRFRHLEPGLNLIPLVLRIRYOSLKEIYINPEOSAVTLNDV 97
 DB 20 VKIVPQGVQWTERGRTKTLQGLSLVFPMDRIGRKINMEQYDLDPSEVISKDVA 79
 QY 98 TLQIDGVLRLMDPKASYGVDEPEYAVTQLAOTTNSELGKLSXDKVFRERESLNASTI 157
 DB 80 NVTIDAVCFIQVIDAPRAAYEVSNLELAINTMTNIRIVLGSMEIDELTSGORDSINSRL 139
 QY 158 VDAINQADMGICRCLREIKIDIHVPPRYKESMOMOVEERRKRAFTVLESEGTRESALNV 217
 DB 140 LRIYVEATNPWCITKTRTIRIDVRPAELISSMNAOMKAEKTKRAVITLAEGRIOAEITLK 199
 QY 218 AEGKQAOILASEAKAEQINO-A-AGEASAVLAKAKAKAEAIRILAAALTOHNGD-AAAS 275
 DB 200 AEGEQSOILAKAEGRQSAFLQAEARERS-----AEAEARAKRMSEALTA--SGDIQAVN 252
 QY 276 LTVAEQVYSAFSKLLAKDSNT--ILLPSNGDVTSMVAQMGV 315
 DB 253 YFVAQKYTEALQOIGSSSNSKVMPL--EASSILMGSIAGI 291
 RESULT 4
 Y692_METTH STANDARD; PRT; 318 AA.
 ID YE20_METTH
 AC 026788;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein MTH692.
 GN MTH692.
 OS Methanobacterium thermoautotrophicum.

OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
 OC Methanobacter.
 OC NCBI_TaxID=145262;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DELTA H;
 RX MEDLINE=98037514; Pubmed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lunn W., Pothier B., Qiu D.,
 RA Spadafora R., Vitare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
 RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 RL J. Bacteriol. 179:7135-7155(1997).
 CC -1- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC
 CC EMBL; AE000848; AAB85197.1; -
 DR InterPro; IPR001107; Band_7.
 DR InterPro; IPR001972; Stomatin.
 DR Pfam; PF01145; Band_7; 1.
 DR PRINTS; PR00721; STOMATIN.
 DR SMART; SM00244; PHB; 1.
 DR PROSITE; PS01270; BAND_7; 1.
 DR Hypothetical protein; Transmembrane; Complete proteome.
 KW TRANSMEM 2 22 POTENTIAL.
 FT SEQUENCE 318 AA: 35425 MW; 4A72C0AC8E99278D CRC64;
 SQ
 Query Match 25.9%; Score 457; DB 1; Length 318;
 Best Local Similarity 36.5%; Pred. No. 6.6e-23;
 Matches 108; Conservative 61; Mismatches 107; Indels 20; Gaps 7;
 QY 45 EAWVERMGRFRHLEPGLNLIPLVLRIRYOSLKEIYINPEOSAVTLNDVTLQIDGV 104
 DB 26 EKGVERLCKYQRYVESGLVITLPEIAIKV-DMREQVVDVPPQEVITKDNVTVVVDV 84
 QY 105 LYLRIMDPKASYGVDEPEYAVTQLAOTTNSELGKLSXDKVFRERESLNASTIADINOA 164
 DB 85 IFYEYVDPEPNNAVYVNDVFOATKTLAQTNLNIGLELDQTLTSREMIQTQLREVLEDA 144
 QY 165 ADCWGRICRLRYEIKIDIHVPPRYKESMOMOVEERRKRAFTVLESEGTRESALNV 224
 DB 145 TDKWGRVRYVRIQRIEPPGDIVEANSSKQKAEKRMRAILAEAGYKQSEIRAEEDKQA 204
 QY 225 QILASAEKAEQINOAGASAVLAKAKAKAEAIRILAAALTOHNGDAAASLTVAPOYS 284
 DB 205 AILAEAG-KAEAIKVA-DANKYRELAINEGQAKAILSFRAHHEDPINDI-IALKYLE 261
 QY 285 AFSKLAKDSNT-ILLPSNGDVTSMVAQMGVYALTKAPVGPDPDLSGSSGRDV 339
 DB 262 ALEKVDGRATKILLP-----VEATGILGS-----IAGISMLSDPDKGV 302
 RESULT 5
 YE20_ARCFU STANDARD; PRT; 249 AA.
 ID YE20_ARCFU
 AC 028852;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein AFl420.
 GN AFl420.

OS Archaeoglobus fulgidus.
 CC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
 CC Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A. 4304 / ATCC 49558;
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kervatage A.R., Graham D.E., Kyriides N.C.,
 RA Fleischman R.D., Kerkvater J., Lee N.H., Sutton G.G., Gill S.,
 RA Kitznes E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodex A., Zhou L.,
 RA Colton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujil C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 CC -1- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AE001006; AAB89829.1; -
 CC TIGR: AF1420; -
 DR InterPro: IPR001107; Band_7.
 DR InterPro: IPR001972; Stomatin.
 DR Pfam: PF01145; Band_7; 1.
 DR PRINTS: PR00721; STOMATIN.
 DR SMART: SM00244; PHB; 1.
 DR PROSITE: PS01270; BAND_7; 1.
 DR Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 3 23 POTENTIAL.
 FT SEQUENCE 249 AA; 28346 MW; 4AFEND58AF6F230 CAC64;
 SO
 Query Match 20.5%; Score 362; DB 1; Length 249;
 Best Local Similarity 36.0%; Pred. No. 8.5e-17;
 Matches 81; Conservative 50; Mismatches 76; Indels 18; Gaps 5;
 Oy 37 VVLF-----VPOEAMVVERMGRRHRLPEGLNLLIPVLDRIYVQSLKEIVINPEQ 89
 Db 12 VVFLLSAVRIKVEKRGVIFRLGVLGARGGLFIIPILENM-VVVDLRTVTVDSQ 70
 Oy 90 SAVTLDNTTLDIGVLYLRIMDPYKASYGVEDPEYAVTQAOQTMRSELGKSDKXVRE 149
 Db 71 EVYTDNTVYKAVNAVYRVVPAKAVTEVPYQYATAQAOQTIRSLIGQALDEVLSE 130
 Oy 150 RESLNASTVDAINQADWCIGRLREIKDIHVPRVKSQMQVQAEERKRAVLESEGR 209
 Db 131 RDKLVKIQOITIDEETNPMGIVKVTAVEIKDVELPEEMRRIMQAQAEERERSKIIIRAG 190
 Oy 210 TRESINVAEGKKQAOQLASEAKA-----EQINQAGEASAVL 248
 Db 191 EYQAAMKLRFA--ADVLA-OSGAILLRLYQTLNEISAEQNTT 231
 RESULT 6
 ID Y827_METUA STANDARD; PRT; 199 AA.
 AC Q58237;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein M0827.
 GN M0827.
 OS Methanococcus jannaschii.
 CC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
 CC Methanococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A. 2661 / ATCC 43067;
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kervatage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Nguyen D.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Furumann J.L., Nguyen D.,
 RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Colton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Kleink H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT jannaschii.";
 RL Science 273:1058-1073(1996).
 CC -1- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U67526; AAB98826.1; -
 CC TIGR: M0827; -
 DR InterPro: IPR001107; Band_7.
 DR InterPro: IPR001972; Stomatin.
 DR Pfam: PF01145; Band_7; 1.
 DR SMART: SM00244; PHB; 1.
 DR PROSITE: PS01270; BAND_7; 1.
 DR Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 7 27 POTENTIAL.
 FT SEQUENCE 199 AA; 22809 MW; 2545CE07DD54337E CAC64;
 SO
 Query Match 20.1%; Score 354.5; DB 1; Length 199;
 Best Local Similarity 38.5%; Pred. No. 2e-16;
 Matches 67; Conservative 46; Mismatches 60; Indels 1; Gaps 1;
 Oy 38 VLFVPOEAMVVERMGRRHRLPEGLNLLIPVLDRIYVQSLKEIVINPEQSAVTLNV 97
 Db 25 IVIVNQEGLIFRGRVIGKLPKGINITIPFLD-VPVKVDKRTVTDIPQEMITKDNA 83
 Oy 98 TLQIDGVLYLRIMDPYKASYGVEDPEYAVTQAOQTMRSELGKSDKXVRE 157
 Db 84 VYKVDVAVYVRIYDEKMLLEVEDYEVAILINAGTTLAIGSMLEVDLKNKREYNKL 143
 Oy 158 VDAINQADWCIGRLREIKDIHVPRVKSQMQVQAEERKRAVLESEGR 211
 Db 144 LEILDRETDAMGVRIEKEVKEIDPEEDIKNMAQQAERLKRRAILLEAEGER 197
 RESULT 7
 ID YF11_PYRHO STANDARD; PRT; 266 AA.
 AC O59180;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein PH1511.
 GN PH1511.
 OS Pyrococcus horikoshii.
 CC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
 OX NCBI_TaxID=53953;
 RN [1]

```

RP SEQUENCE FROM N.A.
RC STRAIN-OT3;
RX MEDLINE-98344137; PubMed-9679194;
RA Kawarabayasi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohtuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
RL DNA Res. 5:55-76(1998).
CC -1- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AP000006; BAA30619.1; -
CC InterPro: IPR001107; Band_7.
CC InterPro: IPR001972; Stomatlin.
CC Pfam: PF01145; Band_7; 1.
CC PRINTS: PR00721; STOMATIN.
CC SMART: SM00244; PHB; 1.
CC PROSITE: PS01270; BAND_7; 1.
CC Hypothetical protein; Transmembrane; Complete proteome.
KM TRANSMEM 7
FT SEQUENCE 266 AA; 29999 MW; 21D06F199C31021F CRC64;
SQ

```

Query Match 18.5%; Score 327.5; DB 1; Length 266;
 Best Local Similarity 33.0%; Pred. No. 1.5e-14;
 Matches 72; Conservative 55; Mismatches 82; Indels 9; Gaps 3;

```

OY 36 TVVLVPOOEAWVVRMGFRHLEPGILNIPVDRIRYVSLKEIVNVEQSATVD 95
DB 23 SAIKIVKEERAVIFRLGVRVAGRGGLFIIPIFEK-AVIYDLRQVLDVYQETITD 81
OY 96 NVTLQIDGVLYRIMDPKASYGVDEPEYAVTQLAQTMRSELGSLKDXFRERESLNA 155
DB 82 NVPVRNAVYFRVVDVPAVAVTQVKNYIMATQSITSLRSYVIGQHLDELLESRDKLNM 141
OY 156 STVDAINQADWCGRCLRYEIKDIIHPPRVKESMQVQAEARRKRAYLSESGTRESAI 215
DB 142 QLQRIIDETDPMGIVKAVTEIKDELPLPGMKAMARQAEERERRRITTLAEARQAAE 201
OY 216 NVAEGKKOAOI-----LASEAKAEQINOAGASAVL 248
DB 202 KUREA---AEIIEHPMALQLRTLOTISDVADKSNVI 236

```

RESULT 8
 ID VJ33_PYRAB STANDARD; PRT; 268 AA.
 AC 09VOY1;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein PAB1933.
 GN PAB1933.
 OS Pyrococcus abyssi.
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
 OX NCBI_TaxID=29292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ORSAY;
 RA Hellig R.;
 RT "Pyrococcus abyssi genome sequence: Insights into archaeal chromosome
 RT structure and evolution.";

```

RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AJ248285; CAB9571.1; -
CC InterPro: IPR001107; Band_7.
CC InterPro: IPR001972; Stomatlin.
CC Pfam: PF01145; Band_7; 1.
CC PRINTS: PR00721; STOMATIN.
CC SMART: SM00244; PHB; 1.
CC PROSITE: PS01270; BAND_7; 1.
CC Hypothetical protein; Transmembrane; Complete proteome.
KM TRANSMEM 1
FT SEQUENCE 268 AA; 30241 MW; AELFBEBBD48A4F84 CRC64;
SQ

```

Query Match 18.2%; Score 321.5; DB 1; Length 268;
 Best Local Similarity 31.9%; Pred. No. 3.8e-14;
 Matches 76; Conservative 54; Mismatches 83; Indels 25; Gaps 5;

```

OY 32 LPRN-----TVVL-----TVPOEAWVVRMGFRHLEPGILNIPVDRIRY 75
DB 3 LPRNFVTITLFIIFLITASAIKIVKEERAVIFRLGVRVAGRGGLFIIPIFEK-AV 61
OY 76 VOSLKEIVNVEQSATVDNVTLOIDGVLYRIMDPKASYGVDEPEYAVTQLAQTMR 135
DB 62 IYDLRQVLDVYQETITDNVPRNAVYFRVVDVPAVAVTQVKNYIMATQSITSLR 121
OY 136 SELGSLKDXFRERESLNAISTVDAINQADWCGRCLRYEIKDIIHPPRVKESMQVQ 195
DB 122 SYVQAHDELLESRDKLNMQLOIRIIDEATDPMGIVKAVTEIKDELPLPGMKAMARQ 181
OY 196 AERRKRAYLSESGTRESAINVAEGKKOAOI-----LASEAKAEQINOAGASAVL 248
DB 182 AERRRARRITTLAEARQAAEKUREA---AEIIEHPMALQLRTLOTISDVADKSNVI 236

```

RESULT 9
 ID BAN7_HUMAN STANDARD; PRT; 287 AA.
 AC P27105; Q14087; Q15609;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Erythrocyte band 7 integral membrane protein (Stomatlin) (Protein
 DE 7.2b).
 GN EPR72 OR BND7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE-91355220; PubMed-1883838;
 RA Hiebl-Dierschmid C.M., Entler B., Glotzmann C., Maurer-Fogy I.,
 RA Stratowa C., Prohaska R.;
 RT "Cloning and nucleotide sequence of cDNA encoding human erythrocyte
 RT band 7 integral membrane protein.";
 RL Biochim. Biophys. Acta 1090:123-124(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-96423038; PubMed-8825639;
 RA Unfried I., Entler B., Prohaska R.;
 RT "The organization of the gene (EPR72) encoding the human erythrocyte
 RT band 7 integral membrane protein (protein 7.2b).";
 RL Genomics 30:521-528(1995).

```

RN [3] SEQUENCE FROM N.A.
RP TISSUE-Bone marrow;
RC Stewart A.C., Hepworth-Jones B.J., Keen J.N., Dash B.J.C.,
RA Argent A.C., Casimir C.M.;
RL Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=96064711; PubMed=7592848;
RA Gallagher P.G., Forget B.G.;
RT "Structure, organization, and expression of the human band 7.2b gene,
a candidate gene for hereditary hydrocytosis.";
RL J. Biol. Chem. 270:26358-26363(1995).
RN [5]
RP SEQUENCE OF 4-24, AND PHOSPHORYLATION SITE.
RX MEDLINE=93385136; PubMed=8373790;
RA Salzer U., Aorn H., Prohaska R.;
RT "Identification of the phosphorylation site on human erythrocyte band
7 integral membrane protein: implications for a monotypic protein
structure.";
RL Blochlm. Biophys. Acta 1151:149-152(1993).
RN [6]
RP PALMITOYLATION.
RX MEDLINE=99268434; PubMed=10338112;
RA Snyers L., Umlauf E., Prohaska R.;
RT "Cysteine 29 is the major palmitoylation site on stomatin.";
RL FEBS Lett. 449:101-104(1999)
CC -1- FUNCTION: THOUGH TO REGULATE CATION CONDUCTANCE.
CC -1- SUBCELLULAR LOCATION: EXPOSED ON THE CYTOPLASMIC SURFACE OF THE
CC MEMBRANE.
CC -1- DISEASE: DEFECTS IN EPB72 IN RED CELLS OF PATIENTS WITH HEREDITARY
CC STOMATOCYTOSIS OR CRYOHYDROCYTOSIS RESULT IN AN INCREASED NA+/K+-
CC PERMEABILITY AND HENCE TO A DISORDER OF CELL VOLUME CONTROL.
CC -1- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X60067; CAA42671.1; -.
DR EMBL; X85116; CAA59436.1; -.
DR EMBL; X85117; CAA59436.1; JOINED.
DR EMBL; M81635; AAS58432.1; -.
DR EMBL; U33931; AAC50296.1; ALT-SEQ.
DR EMBL; U33925; AAC50296.1; JOINED.
DR EMBL; U33925; AAC50296.1; JOINED.
DR EMBL; U33927; AAC50296.1; JOINED.
DR EMBL; U33928; AAC50296.1; JOINED.
DR EMBL; U33929; AAC50296.1; JOINED.
DR EMBL; U33930; AAC50296.1; JOINED.
DR PIR; S17659; S17659.
DR MIM; 185000; -.
DR MIM; 185000; -.
DR InterPro: IPR001107; Band_7.
DR InterPro: IPR001972; Stomatin.
DR Pfam: PF01145; Band_7; 1.
DR PRINTS; PR00721; STOMATIN.
DR SMART; SM00244; PHB; 1.
DR PROSITE; PS01270; BAND_7; 1.
KW Erythrocyte; Transmembrane; Phosphorylation; Lipoprotein; Palmitate.
FT INIT_MET 0
FT TRANSMEM 0
FT DOMAIN 25 53 POTENTIAL.
FT MOD_RES 54 287 CYTOPLASMIC (POTENTIAL).
FT LIPID 9 9 PHOSPHORYLATION (BY CAPK).
FT LIPID 29 29 PALMITATE.
FT LIPID 86 86 PALMITATE (PARTIAL).
FT CONFLICT 5 5 H -> D (IN REF. 3)
SQ SEQUENCE 287 AA: 31599 MW: 4852FE286ECC337 CRC64:

```

[illegible]

UNCL_CABEEL	STANDARD:	PRT:	282 AA.
ID UNCL_CABEEL			
AC 021190:			
DT 01-NOV-1997 (Rel. 35, Created)			
DT 01-NOV-1997 (Rel. 35, Last sequence update)			
DT 15-DEC-1998 (Rel. 37, Last annotation update)			
DE UNC-1 protein.			
GN UNC-1 OR K0356.5.			
OS Caenorhabditis elegans.			
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;			
OC Rhabditiidae; Peloderinae; Caenorhabditis.			
OX NCBI_TaxID=62319;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN-BRISTOL N2.			
RA Latreille P., Gattung S.;			
RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.			
CC -1- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.			
CC -----			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC the European Bioinformatics Institute. There are no restrictions on its			
CC use by non-profit institutions as long as its content is in no way			
CC modified and this statement is not removed. Usage by and for commercial			
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC or send an email to license@isb-sib.ch).			
CC -----			
CC DR EMBL: U55375; AAC69044.1; "			
DR WormRep: K0356.5; CE07331.			
DR InterPro: IPR001107; Band.7.			
DR InterPro: IPR001972; Stomatlin.			
DR Pfam: PF01145; Band.7; 1.			
DR PRINTS: PR00721; STOMATLIN.			
DR SMART: SM00244; PHB: 1.			
DR PROSITE: PS01270; BAND.7; 1.			
DR Transmembrane; Multigene family.			
KW TRANSMEM 24			
FT POTENTIAL. 44			
SQ SEQUENCE 282 AA: 31246 MW: 7855DF6C02268480 CRC64:			
Query Match 15.7%; Score 277.5; DB 1; Length 282;			
Best Local Similarity 30.2%; Pred. No.2,8e-11;			
Matches 64; Conservative 51; Mismatches 92; Indels 5; Gaps 3;			
QY 41 VQEQAWVVERNGR-FHRILTEPGNILLPLVDLRIRYVOSLKEIVINPEQSAVTLDNVT 98			
Db 55 IKEYEYVIFRIGRLVEFGARGGMFIIPICIDYTKI-DLRVSYVAAPDEILSKDSVT 113			
QY 99 LQIDGVLYLRINDPFRKASGVDEPEAVYQLAQTMRSLGSLSKDKYVREESLNASTV 158			

DB 114 VSDVAVYFPTSPDIASVNNVDIAISTKLLAOTTLRNALGKLTLEMLTTERAINOLCE 173
 QY 159 DAINOADCWICRLREYKIDIHVPRVKSOMQVAEERKRATVLESEGTRE--AIN 216
 DB 174 TILDEGTEHMGVVEREVNDIRLPOLUTRAMAEAEAKRAKVAAEGEOKASALK 233
 QY 217 VAEGRKQAOILASEAEKAEQINDAGEASAVL 248
 DB 234 EAADVIOANFVALQLRHLQALNSIAAEHNSTI 265

RESULT 11
 BAN7_MOUSE
 ID BAN7_MOUSE STANDARD: PRT: 284 AA.
 AC P54116; Q60744; 062455;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Erythrocyte band 7 Integral membrane protein (Stomatlin) (Protein 7.2b).
 GN EPB72 OR EPB7.2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN 111
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C; TISSUE-Bone marrow;
 RX MEDLINE-97080556; PubMed-8921901;
 RA Schlegel W., Unfried I., Prohaska R.;
 RT "Cloning and analysis of a cDNA encoding the BALB/c murine erythrocyte band 7 integral membrane protein.";
 RL Gene 178:115-118(1996).
 RN 121
 RP SEQUENCE FROM N.A.
 RC STRAIN-C57BL/6J;
 RA Gallagher P.G., Turetsky T., Mentzer W.C.;
 RT Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
 RN 313
 RP SEQUENCE FROM N.A.
 RX MEDLINE-96374834; PubMed-8786142;
 RA Gallagher P.G., Turetsky T., Mentzer W.C.;
 RT "Genomic organization and 5'-flanking DNA sequence of the murine stomatin gene (Epb72).";
 RL Genomics 34:410-412(1996).
 CC 1- FUNCTION: THOUGHT TO REGULATE CATION CONDUCTANCE.
 CC 1- SUBCELLULAR LOCATION: EXPOSED ON THE CYTOPLASMIC SURFACE OF THE MEMBRANE.
 CC 1- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X91043; CAA62503.1; -
 DR EMBL: U17297; AAB15024.1; -
 DR EMBL: U50999; AAB18857.1; ALT_SEQ.
 DR EMBL: U50993; AAB18857.1; JOINED.
 DR EMBL: U50994; AAB18857.1; JOINED.
 DR EMBL: U50995; AAB18857.1; JOINED.
 DR EMBL: U50996; AAB18857.1; JOINED.
 DR EMBL: U50997; AAB18857.1; JOINED.
 DR EMBL: U50998; AAB18857.1; JOINED.
 DR MGD: MGI:95403; EPB7.2.
 DR InterPro: IPR001107; Band_7.
 DR InterPro: IPR001972; Stomatlin.
 DR Pfam: PF01145; Band_7; 1.
 DR PRINTS: PR00721; STOMATIN.
 DR SMART: SM00244; PHB; 1.
 DR SMART: SM00244; PHB; 1.

DR PROSITE; PS01270; BAND_7; 1.
 KW Erythrocyte; Transmembrane; PHOSPHORYLATION; Lipoprotein; Palmitate.
 FT TRANSMEM 32 52
 FT DOMAIN 53 284
 FT LIPID 30 30
 FT LIPID 87 87
 FT CONFLICT 37 37
 FT CONFLICT 40 40
 FT CONFLICT 43 43
 FT CONFLICT 91 91
 FT CONFLICT 273 273
 FT CONFLICT 283 283
 SQ SEQUENCE 284 AA; 31403 MW; 8E49276511ACB84F CRC64;

Query Match 15.7%; Score 277.5; DB 1; Length 284;
 Best Local Similarity 30.7%; Pred. No. 2.8e-11;
 Matches 63; Conservative 50; Mismatches 79; Indels 13; Gaps 4;

QY 38 VLEVPQGEAWVERMGRFRHILE----PGLNLIPIVLDIRIRYQSLKEIVINVEQSAV 92
 DB 54 IKIVKEVERVIIFRLG---RLLOGGAKGPGFLFPLCTDSFIYV-DKRTISFDIPQEV 109
 QY 93 TLDNVTLDIGVLYLRIMDPYKASYGVDEPEYAVTOLAQTMTSELKLSXKVFRRER 152
 DB 110 TKDSVTISVDGVVYRYQNAITNADSAITRLAQTTLRNALGKLTLEMLTTERAINOLCE 169
 QY 153 LMSIYVAINOADCWICRLREYKIDIHVPRVKSOMQVAEERKRATVLESEGTRE 212
 DB 170 IAHMSTLDDATDDWIKERVEIKDKLPVOLQRAMAEAEAKRAKVAAEGEOKASALK 226
 QY 213 SAINVAEGRKQAOILASEAEKAEQI 237
 DB 227 -EMNASRALKEASVITESPALQL 250

RESULT 12
 ST02_CAEEL
 ID ST02_CAEEL STANDARD: PRT: 280 AA.
 AC 019958;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE ST0-2 protein.
 GN ST0-2 OR F32A6.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OX Rhabditidae; Pelodermidae; Caenorhabditis.
 OX NCBI_TaxId=6239;
 RN 111
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Wu X.;
 RT Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 CC 1- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U40409; AAB1390.1; -
 DR WormPep: F32A6.5; CE04473.
 DR InterPro: IPR001107; Band_7.
 DR InterPro: IPR001972; Stomatlin.
 DR Pfam: PF01145; Band_7; 1.
 DR PRINTS: PR00721; STOMATIN.
 DR SMART: SM00244; PHB; 1.
 DR PROSITE; PS01270; BAND_7; 1.
 DR Transmembrane; Multigene family.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 18:20:57 ; Search time 63.67 Seconds

(without alignments)
896.843 Million cell updates/sec

Title: us-09-898-216-1

Perfect score: 1767
Sequence: 1 MLARAARGHWGPPAEGLSTG.....RDVQGTDAKXDELDKVKMS 356

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mmc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_virus:*

16: sp_bacteriaph:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1646	93.2	356	4 Q9UJZ1	Q9UJZ1 homo sapien
2	1639	92.8	356	4 Q96FY2	Q96FY2 homo sapien
3	1613	91.3	342	4 Q9P042	Q9P042 homo sapien
4	1605	90.8	357	4 Q60376	Q60376 homo sapien
5	1537.5	87.0	353	11 Q99JB2	Q99JB2 mus musculu
6	1526.5	86.4	353	11 Q9DCG8	Q9DCG8 mus musculu
7	1064.5	60.2	323	5 Q9WLF7	Q9WLF7 drosophila
8	957.5	54.2	334	5 Q9XVP9	Q9XVP9 caenorhabdi
9	861.5	48.8	354	3 Q60121	Q60121 schizosacch
10	826	46.7	411	10 Q93VP9	Q93VP9 arabidopsis
11	826	46.7	515	10 Q9T082	Q9T082 arabidopsis
12	816.5	46.2	401	10 Q9LW00	Q9LW00 arabidopsis
13	811.5	45.9	312	16 Q921G8	Q921G8 rickettsia
14	806	45.6	311	16 Q9ZDK0	Q9ZDK0 rickettsia
15	794.5	45.0	394	10 Q9MS85	Q9MS85 zea mays (m
16	722	40.9	374	5 Q77376	Q77376 plasmodium

17	716	40.5	315	16 Q9JR80	Q9JR80 neisseria m
18	714.5	40.4	315	2 Q9JPH5	Q9JPH5 neisseria m
19	714	40.4	315	2 Q9JPN3	Q9JPN3 neisseria m
20	712.5	40.3	315	2 Q9JPC5	Q9JPC5 neisseria m
21	712.5	40.3	315	16 Q9JRP6	Q9JRP6 neisseria m
22	711.5	40.3	315	2 Q9K2T7	Q9K2T7 neisseria g
23	711.5	40.3	315	2 Q9JPH3	Q9JPH3 neisseria m
24	711.5	40.3	315	2 Q9JPN2	Q9JPN2 neisseria m
25	664	37.6	357	5 Q9U4M5	Q9U4M5 leishmania
26	497	28.1	344	16 Q9RS11	Q9RS11 delinococcus
27	496.5	28.1	337	16 Q9PGK3	Q9PGK3 xylolla fas
28	486	27.5	310	16 Q9A331	Q9A331 caudobacter
29	477	27.0	306	2 Q9AH28	Q9AH28 photorhabd
30	475	26.9	319	2 Q9X9Z6	Q9X9Z6 streptomyce
31	474.5	26.9	325	2 Q9F507	Q9F507 escherichia
32	473	26.8	307	16 Q9CME1	Q9CME1 pasteurilla
33	467.5	26.5	306	16 Q9KTC6	Q9KTC6 vibrio chol
34	464.5	26.3	343	2 Q9K458	Q9K458 streptomyce
35	460	26.0	298	17 Q98205	Q98205 pyrococcus
36	460	26.0	299	17 Q9UYE4	Q9UYE4 pyrococcus
37	454	25.7	328	16 Q9ZM51	Q9ZM51 rhizobium l
38	444.5	25.2	374	16 Q9CBM5	Q9CBM5 mycobacteri
39	442.5	25.0	322	16 Q9TK67	Q9TK67 clostridium
40	442	25.0	320	2 Q9R005	Q9R005 clostridium
41	430	24.3	316	16 Q98H59	Q98H59 rhizobium l
42	423.5	24.0	305	16 Q9WZK0	Q9WZK0 thermotoga
43	417.5	23.6	392	17 Q9HSA8	Q9HSA8 halopotaeri
44	412	23.3	309	16 Q98RE1	Q98RE1 mycoplasma
45	376.5	21.3	253	16 Q67057	Q67057 aquilex aeo

ALIGNMENTS

RESULT 1

Q9UJZ1 1 PRELIMINARY: PRT; 356 AA.

AC Q9UJZ1; 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)

DE MEMBRANE ASSOCIATED PROTEIN SLP-2 (STOMATIN-LIKE PROTEIN 2) (STOMATIN-LIKE 2) (CDNA FLJ14499 FIS, CLONE NT2RMI00080, WEAKLY SIMILAR TO UNC-1 PROTEIN) (STOMATIN-LIKE 2).

GN HUSLP2 OR SLP2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OC NCBI_TaxID=9606;

RP [1]

RP SEQUENCE FROM N.A.

RP TISSUE=HEART MUSCLE;

RA Wang Y., Morrow J.S.;

RT "Identification and Characterization of Human SLP-2, a Novel Homologue of Stomatlin (band 7.2b) Present in Erythrocytes and Other Tissues.";

RT Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.

RN [2]

RN SEQUENCE FROM N.A.

RN TISSUE=BRAIN;

RC Owczaek C.M., Treutlein H.R., Portbury K.J., Gulluyan L.M., Kola I., Hertzog P.J.;

RA "A novel member of the stomatin/EPB72/MC-2 family, stomatin-like protein 2 (SLP2), is ubiquitously expressed and localizes to HSA chromosome 9p13.1.";

RT Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.

RL [3]

RL SEQUENCE FROM N.A.

RL TISSUE=LUNG CARCINOMA, SKIN, AND MELANOTIC MELANOMA.;

RA Strusberg R.;

RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.

RN [4]

RN SEQUENCE FROM N.A.

RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,

RA Nishikawa T., Nagai K., Sugano S., Shitatori A., Sudo H.,
RA Megatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakami K., Ono Y., Takiguchi S.,
RA Matsunobu S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Makematsu A., Nakamura Y., Nagahari K., Masuno Y.,
RA Niinomiya K., Iwyanagi T.,
RT "NEDO human cDNA sequencing project."
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE-PANCREAS, EPITHELIOID CARCINOMA;
RA Strausberg R.;
RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF190167; AAF09142.1; -;
DR EMBL: AF282596; AAF91466.1; -;
DR EMBL: BC003025; AAH03025.1; -;
DR EMBL: BC002442; AAH02442.1; -;
DR EMBL: AK027405; BAB55091.1; -;
DR EMBL: BC014990; AAH14990.1; -;
DR InterPro: IPR001107; Band_7.
DR Pfam: PF01145; Band_7; 1.
DR InterPro: IPR001972; Stomatin.
DR PRINTS: PR00721; STOMATIN.
DR SMART: SM00244; PHB; 1.
SQ SEQUENCE 356 AA; 38534 MW; 672331B57C82654E CRC64;

Query Match 93.2%; Score 1646; DB 4; Length 356;
Best Local Similarity 95.3%; Pred. No. 1,2e-97;
Matches 341; Conservative 2; Mismatches 11; Indels 4; Gaps 2;

OY 1 MLARAARGHMGPFPAEG--LSTGFMPRSGRASSGGLPRNTVLFVPOQEAAMVVERMGFRHRI 58
DB 1 MLARAARGTGALLRGLSLASGRAPR--RASSGGLPRNTVLFVPOQEAAMVVERMGFRHRI 58
OY 59 LEPLGNTLIPVLDRIYVQSLKEIYINPEQSAVTLDNVTQIDGVLYLRIMDPKASYG 118
DB 59 LEPLGNTLIPVLDRIYVQSLKEIYINPEQSAVTLDNVTQIDGVLYLRIMDPKASYG 118
OY 119 VEDPEYAVTQLAQTMRSELKLSXDKYFRERRESINASTIVAINOACDCKIRCLRYEIK 178
DB 119 VEDPEYAVTQLAQTMRSELKLSXDKYFRERRESINASTIVAINOACDCKIRCLRYEIK 178
OY 179 DIHVPVRKESMOMQVEAEERRKRATVLESEGTRESAINVAEGKKQAQITLASEAEKAEQIN 238
DB 179 DIHVPVRKESMOMQVEAEERRKRATVLESEGTRESAINVAEGKKQAQITLASEAEKAEQIN 238
OY 239 QAAGEASAVLAKAKAKAIAIRILAAALTOHNGDAAASTLVAEQYVSARSKLAKDSNTILL 298
DB 239 QAAGEASAVLAKAKAKAIAIRILAAALTOHNGDAAASTLVAEQYVSARSKLAKDSNTILL 298
OY 299 PSNPGDVTSMVAQAMGVGALTTPKAPVPTPSLSGSSRDVQGTDPASXDELDRYKMS 356
DB 299 PSNPGDVTSMVAQAMGVGALTTPKAPVPTPSLSGSSRDVQGTDPASXDELDRYKMS 356
RESULT 2
O96FY2 PRELIMINARY; PRT; 356 AA.
AC O96FY2
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE STOMATIN-LIKE 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG CARCINOMA;
RA Strausberg R.;
RL Submitted (Jul-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: BC010152; AAH10152.1; -;
SQ SEQUENCE 356 AA; 38518 MW; 6FC0BF899E1919A7 CRC64;
Query Match 92.8%; Score 1639; DB 4; Length 356;
Best Local Similarity 95.0%; Pred. No. 3,4e-97;
Matches 340; Conservative 2; Mismatches 12; Indels 4; Gaps 2;

OY 1 MLARAARGHMGPFPAEG--LSTGFMPRSGRASSGGLPRNTVLFVPOQEAAMVVERMGFRHRI 58
DB 1 MLARAARGTGALLRGLSLASGRAPR--RASSGGLPRNTVLFVPOQEAAMVVERMGFRHRI 58
OY 59 LEPLGNTLIPVLDRIYVQSLKEIYINPEQSAVTLDNVTQIDGVLYLRIMDPKASYG 118
DB 59 LEPLGNTLIPVLDRIYVQSLKEIYINPEQSAVTLDNVTQIDGVLYLRIMDPKASYG 118
OY 119 VEDPEYAVTQLAQTMRSELKLSXDKYFRERRESINASTIVAINOACDCKIRCLRYEIK 178
DB 119 VEDPEYAVTQLAQTMRSELKLSXDKYFRERRESINASTIVAINOACDCKIRCLRYEIK 178
OY 179 DIHVPVRKESMOMQVEAEERRKRATVLESEGTRESAINVAEGKKQAQITLASEAEKAEQIN 238
DB 179 DIHVPVRKESMOMQVEAEERRKRATVLESEGTRESAINVAEGKKQAQITLASEAEKAEQIN 238
OY 239 QAAGEASAVLAKAKAKAIAIRILAAALTOHNGDAAASTLVAEQYVSARSKLAKDSNTILL 298
DB 239 QAAGEASAVLAKAKAKAIAIRILAAALTOHNGDAAASTLVAEQYVSARSKLAKDSNTILL 298
OY 299 PSNPGDVTSMVAQAMGVGALTTPKAPVPTPSLSGSSRDVQGTDPASXDELDRYKMS 356
DB 299 PSNPGDVTSMVAQAMGVGALTTPKAPVPTPSLSGSSRDVQGTDPASXDELDRYKMS 356

RESULT 3
O9P042 PRELIMINARY; PRT; 342 AA.
AC O9P042
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE STOMATIN-LIKE 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RX MEDLINE=20499367; PubMed=11042152;
RA Zhang Q.H., Ye M., Wu X.Y., Ren S.X., Zhao M., Zhao C.J., Fu G.,
RA Shen Y., Fan H.Y., Lu G., Zhong M., Xu X.R., Han Z.G., Zhang J.W.,
RA Tao J., Huang Q.H., Zhou J., Hu G.X., Gu J., Chen S.J., Chen Z.J.;
RT "Cloning and functional analysis of cDNAs with open reading frames for
300 previously undefined genes expressed in CD34+ hematopoietic
stem/progenitor cells."
RL Genome Res. 10:1546-1560(2000).
DR EMBL: AF161458; AAF29073.1; -;
DR InterPro: IPR001107; Band_7.
DR InterPro: IPR001972; Stomatin.
DR Pfam: PF01145; Band_7; 1.
DR PRINTS: PR00721; STOMATIN.
DR SMART: SM00244; PHB; 1.
SQ SEQUENCE 342 AA; 37145 MW; 7958C0E3BDBE53E5 CRC64;

Query Match 91.3%; Score 1613.5; DB 4; Length 342;
Best Local Similarity 97.1%; Pred. No. 1,4e-95;
Matches 330; Conservative 3; Mismatches 4; Indels 3; Gaps 1;
OY 17 LSTGFMPRSGRASSGGLPRNTVLFVPOQEAAMVVERMGFRHRILEPLGNTLIPVLDRIYV 76
DB 6 LASG---RACRASSGGLPRNTVLFVPOQEAAMVVERMGFRHRILEPLGNTLIPVLDRIYV 62

299 PSNSDVTSMVAQAMGVYALTKAPVPCANQ--SSQSRDVOATDTSI--EELGVRKLS 353


```

QY 39 LEVPOEAMVVERMGFRHFRILEPGLNILLIPVLDRIYVOSLKEIYINVEPOSATVLDNT 98
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MEVPOQEAAMVVERMGFRHFRILDPGLNILLIPVADKIKYVOSLKEIADIVPKOSAITSDNT 60
QY 99 LOIDGVLYLRIMDPKASYGVDEPEYATVOLAQTMRSELGKLSXDKYFRRESLNASTV 158
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 LSIDVLYLRITDPKASYGVDEPEFAITQLAQTMRSELGKMSDKYFRRESLNASTV 120
QY 159 DAINOACWGCIRCLRYEIKOIHPVRKESOMQVEARRRATVLESEGRRESAINVA 218
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 DSINKASEAMGIACIRRYEIRDIRLPTFRHEAMQVVEARRRATVLESEGREAEINIA 180
QY 219 EGKROAOILASEAEKAEQINOAGEASAVLAKAKAEAIRILAAALTOHNGDAASLTV 278
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 EGKRSRILASEAEQEHINKASGEAAIIVADARSLAIKASLSHLDQGNASLTU 240
QY 279 AEQYVASAKLAKDSNTILLPSNPGDVTSMVAQMGVYALTKAPPGTSDLSGSSGSD 338
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 AEQYGAERKLAKTNTMLPSNPGDVNGFVAQALAVYVHVSNS-----NOATRKSEN 293
QY 339 VOGTDA-----SXDELDRIYKMS 356
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 294 VKGVACILMAKSVEYKELOEDKSSVKMN 321

RESULT 8
O9XVP9 PRELIMINARY: PRT: 334 AA.
ID O9XVP9
AC O9XVP9
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE F30A10.5 PROTEIN.
GN F30A10.5
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RA SEQUENCE FROM N.A.
RL Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RP MEDLINE=99069613; PubMed=9851916;
RX none;
RT "Genome sequence of the nematode C. elegans: A platform for
   investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; Z81072; CAB03018.1; -.
DR InterPro: IPR001107; Band 7.
DR InterPro: IPR001972; Stomatin.
DR Pfam; PF01145; Band_7; 1.
DR PRINTS; PRO0721; STOMATIN.
DR SMART; SM00244; PHB; 1.
SQ SEQUENCE 334 AA; 36704 MW; 10F98BB9A44E5ED CRC64;

Query Match 54.2%; Score 957.5; DB 5; Length 334;
Best Local Similarity 61.2%; Pred. No. 1.1e-53;
Matches 188; Conservative 58; Mismatches 46; Indels 15; Gaps 2;
QY 29 SSGLP-----RNTVLFVPOOEAMVVERMGFRHFRILEPGLNILLIPVLDRIYV 76
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 24 SSTPLAVATSSQAAAHNTVINEVPOOEAMVVERMGFRHFRILEPGLNILLIPVLDRIYV 83
QY 77 OSLKEIYINVEPOSATVLDNTVLAQIDGVLYLRIMDPK---ASYGEDEPEYATVOLAQT 133
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 84 QNLRELAIEIPGGAITTDNVQLRLDGVLYLRVFPDYKACDASYGVDEPEYATVOLAQT 143
QY 134 MSELGKLSXDKYFRRESLNASTVDAINOAADCMGICIRCLRYEIKDIHVPVRKESOMQ 193
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 144 MRSEVKINLDTVFRERELNENIVFAINKASAPWGICMRYEIRDMQPSKIOEQMOMQ 203

```

```

QY 194 VEAERKRATVLESEGRRESAINVAEGKROAOILASEAEKAEQINOAGEASAVLAKAKA 253
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 204 VEAERKRRAIIESEGRRESAINVAEGKRSKSAIASEAVOAEIRINVAKEAEAVILAKAS 263
QY 254 KAEAIRILAAALTOHNGDAASLTVAEOYVASFSKLAKDSNTILLPSNPGDVTSMVAQAM 313
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 264 RAKAIERIALALEKOGANAAAGLTVAEQYVGAFGNLAKSNTVLPANLSDPGSNVSOAL 323
QY 314 GYVGALT 320
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 324 AVYDSL 330

RESULT 9
O60121 PRELIMINARY: PRT: 354 AA.
ID O60121
AC O60121
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE HYPOTHETICAL 39.3 KDA PROTEIN C1665.07C IN CHROMOSOME II.
GN SPBC1665.07C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_Taxid=4896;
RN [1]
RA SEQUENCE FROM N.A.
RP STRAIN=972;
RC LYNE M., Wood V., Rajandream M.A., Barrell B.G., Beck A.,
   Reinhardt R.;
RL Submitted (May-1998) to the EMBL/Genbank/DBJ databases.
CC -i- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.
DR EMBL; AL023554; CAI19027.1; -.
DR InterPro: IPR001107; Band 7.
DR InterPro: IPR001972; Stomatin.
DR Pfam; PF01145; Band_7; 1.
DR SMART; SM00244; PHB; 1.
DR PROSITE; PS01270; BAND_7; FALSE_NRG.
KW Hypothetical protein.
SQ SEQUENCE 354 AA; 39274 MW; 66B09E6A12BDC030 CRC64;

Query Match 48.8%; Score 861.5; DB 3; Length 354;
Best Local Similarity 56.7%; Pred. No. 1.7e-47;
Matches 170; Conservative 56; Mismatches 71; Indels 3; Gaps 2;
QY 35 NNTVLFVPOOEAMVVERMGFRHFRILEPGLNILLIPVLDRIYVOSLKEIYINVEPOSATV 94
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 50 NNTIKVPOOVAVVERMGFRHFRILTPGVAFLAPITDKIAYINSLERALEIPTQSAITL 109
QY 95 DNVTLQIDGVLYLRIMDPKASYGVDEPEYATVOLAQTMRSELGKLSXDKYFRRESLN 154
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 DNVSLGLDGVLYIQVDPKASYGVDEADYVAISQLAQTMRSEIGRLTDHVLREHQSIN 169
QY 155 ASLVAINOAADCMGICIRCLRYEIKDIHVPVRKESOMQVEARRRATVLESEGRRES 214
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 170 IHTTDAINKAASWGCIRCLRYEIRDIRPPESSVYMAHQOVSARORAEILESEGRQAA 229
QY 215 INVAEGKROAOILASEAEKAEQINOAGEASAVLAKAKAEAIRILAAAL-TOHNGDA 273
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 230 INVAEGDKOAEILDSGCKIKITINSALAEQAIRERASATASIVLAADSIRKQEGLEA 289
QY 274 ASLVAEQYVASFSKLAKDSNTILLPSNPGDVTSMVAQMGVYALTK--APVPGTSD 331
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 290 VSLYIAQYITNGLAKASNSMIVPASTSDVSGVVAQALSLIFKQVSKTAPDKSTPKEL 349

RESULT 10
O93VP9 PRELIMINARY: PRT: 411 AA.
ID O93VP9

```

AC 093VP9;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HYPOTHEORETICAL 45.0 KDA PROTEIN.
 GN ATG27580.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
 RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
 RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamita A., Karlin-Neumann G.,
 RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
 RA Natsume M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
 RA Shinozaki K., Shinozaki A., Tracy S.E., Shinozaki K., Davis R.W.,
 RA Ecker J.R., Theologis A.;
 RT "Full length cDNA of gene AT4g27580 (GI:7269612).";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,
 RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,
 RA Jones T., Kamita A., Karlin-Neumann G., Kawai J., Kim C., Koesema E.,
 RA Lam B., Lin J., Meyers M.C., Miranda M., Natsume M., Nguyen M.,
 RA Palm C.J., Sakurai T., Satou M., Seki M., Shinozaki K., Shinozaki A.,
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RT "Full length cDNA of gene AT4g27580 (GI:7269612).";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY0349109; AAL15215.1; -
 DR EMBL: AY034924; AAK59431.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 411 AA; 45020 MW; 94E24ED4C9B89633 CRC64;

Query Match 46.7%; Score 826; DB 10; Length 411;
 Best Local Similarity 50.1%, Pred. No.3.8e-45;
 Matches 168; Conservative 67; Mismatches 80; Indels 20; Gaps 2;
 Oy 29 SSGLEPRNTVTL-----EYPOGEAMVVERMGRFHRILEPGLNLLIPVLDRIYVOSL 79
 Db 45 SAGYPSNSFOLTPPTNMGIRIVPERKAFVIERFGKYATTLPSGIHFLIPFVDRIYVHSL 104
 Oy 80 KEIYINVPESASVLTNDYNTLOIDGVLVLRIMDPYKASGVDEPEYAVQLOATTMRSELG 139
 Db 105 KEAIPINOTAITKDNVSIHIDGVLVYKIVDPKLASTGVESPIYAVVOLAOATTMRSELG 164
 Oy 140 KLSXDKVFRERESLNASIVDAINQADCGIRCLRYETIKDIIHVPVRESMOMOVEAERR 199
 Db 165 KITLDTKTEERDITNEKIYEAINVNAKDMGLOCLRYEIRIDIMPPIGVRAEMQAEERK 224
 Oy 200 KRAIVLESEGTRESAINVAEGKKAOIILASEAKAEQINQAGESAIVLAKKAAEAIR 259
 Db 225 KRAQILLESSEGEROSHININADGKSSVILASEAKMDVNRAGAEALILARQAAKAGLV 284
 Oy 260 ILAALITQHNDAASLVAEYVSAFSLAKDSNTILLPSNPGVTSVVAQMGVYAL 319
 Db 285 LLSQSLKTEGVEAASLVAEDYITAFGNIAKEGIMILLPSGASNPASIMIAQALTMYSLS 344
 Oy 320 TKAPVPGTPTDSLSSGSSRDVQGTASXDEELDRVK 354
 Db 345 -----VINGRSKDHQETQALDFTDELELE 368

RESULT 11
 Q9T082 PRELIMINARY: PRT: 515 AA.
 AC Q9T082;

DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE HYPOTHEORETICAL 55.9 KDA PROTEIN.
 GN T29A15.70 OR AT4G27580.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Beyan M., Van Der Schueren J., Chuang Y.-J., Voet M., Robben J.,
 RA Volckaert G., Hobeisel J., Mewes H.W., Mayer K.F.X., Schueller C.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Robben J., Grymoprez B., Volckaert G., Mewes H.W., Lemcke K.,
 RA Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL035602; CAB8270.1; -
 DR EMBL: AL161571; CAB81408.1; -
 DR InterPro: IPR001107; Stomatin.
 DR InterPro: IPR001972; Stomatin.
 DR Pfam: PF01145; Band_7; 1.
 DR PRINTS: PR00721; STOMATIN.
 DR SMART: SM00244; PHB; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 515 AA; 55923 MW; 420FE3DB5B7AFDBA CRC64;

Query Match 46.7%; Score 826; DB 10; Length 515;
 Best Local Similarity 50.1%, Pred. No.5.1e-45;
 Matches 168; Conservative 67; Mismatches 80; Indels 20; Gaps 2;
 Oy 29 SSGLEPRNTVTL-----EYPOGEAMVVERMGRFHRILEPGLNLLIPVLDRIYVOSL 79
 Db 45 SAGYPSNSFOLTPPTNMGIRIVPERKAFVIERFGKYATTLPSGIHFLIPFVDRIYVHSL 104
 Oy 80 KEIYINVPESASVLTNDYNTLOIDGVLVLRIMDPYKASGVDEPEYAVQLOATTMRSELG 139
 Db 105 KEAIPINOTAITKDNVSIHIDGVLVYKIVDPKLASTGVESPIYAVVOLAOATTMRSELG 164
 Oy 140 KLSXDKVFRERESLNASIVDAINQADCGIRCLRYETIKDIIHVPVRESMOMOVEAERR 199
 Db 165 KITLDTKTEERDITNEKIYEAINVNAKDMGLOCLRYEIRIDIMPPIGVRAEMQAEERK 224
 Oy 200 KRAIVLESEGTRESAINVAEGKKAOIILASEAKAEQINQAGESAIVLAKKAAEAIR 259
 Db 225 KRAQILLESSEGEROSHININADGKSSVILASEAKMDVNRAGAEALILARQAAKAGLV 284
 Oy 260 ILAALITQHNDAASLVAEYVSAFSLAKDSNTILLPSNPGVTSVVAQMGVYAL 319
 Db 285 LLSQSLKTEGVEAASLVAEDYITAFGNIAKEGIMILLPSGASNPASIMIAQALTMYSLS 344
 Oy 320 TKAPVPGTPTDSLSSGSSRDVQGTASXDEELDRVK 354
 Db 345 -----VINGRSKDHQETQALDFTDELELE 368

RESULT 12
 Q9LVM0 PRELIMINARY: PRT: 401 AA.
 AC Q9LVM0;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, last annotation update)
DE SIMILARITY TO STOMATIN LIKE PROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.,
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:31-63(2000).
DR EMBL: AB018115; BAA97132.1; -
DR InterPro: IPR001107; Band_7.
DR InterPro: IPR001972; Stomatlin.
DR Pfam: PF01145; Band_7; 1.
DR PRINTS: PR00721; STOMATIN.
DR SMART: SM00244; PHB; 1.
SQ SEQUENCE 401 AA; 43607 MW; 28E88D732A370AB6 CRC64;

Query Match 46.2%; Score 816.5; DB 10; Length 401;
Best Local Similarity 55.1%; Pred. No. 1.5e-44;
Matches 161; Conservative 58; Mismatches 72; Indels 1; Gaps 1;

OY 33 PRTVLFVPOQEAAMVVERMGFRHILEPGLNLIIPVDRIRYVOSLKEIYINVPESAV 92
DB 101 PSMWGRIYPERKACIYERFKHTLPACIHFLVFPVDRIAYVHSLEKALYIGQTAI 160
OY 93 TLDNVTLDIGVLYLRIMDPYKASVGEDEPEYAVTOLAQTMRSELGKLSXDKVFRERES 152
DB 161 TDMNVSIHIDGVLYXKIVDPKLASGVENPIYAVMOLAQTMRSELGKITLDKTFEEDRT 220
OY 153 LNASYDAINQADDCGICRLEIKDIHPPRVKESMOMQVAEKRRKATVLESECTRE 212
DB 221 LNEKIYEAINVAKKDWGLCCRLREIRIDIPPNCGVRVAMEQAERKKRAQJLSESEGRQ 280
OY 213 SAIVNAGKKAQOILASEAEKAEQINQAGEASAVLAKKAKAEAIRILAAALTOHNGDA 272
DB 281 AHNTRADGKSSVILESEPAAMDQVNRAGEAAILARAQATRKGLAMVSQSLKEGGE 340
OY 273 AASLVAAEQYVSASFSLKARDSTNTLLPSNPGDVTSMVAQAMGYV-GALTKAP 323
DB 341 AASLVAAEQYIAFGKIAKEGFTMLLPNSVNDPASHMAIQALMGYKGLSTKVP 392
RESULT 13
OY2IG8
ID O92IG8 PRELIMINARY; PRT; 312 AA.
AC O92IG8;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE HYPOTHETICAL PROTEIN RC0452.
GN RC0452.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MALISH 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Nester-Audiffren P., Fournier P.-E., Barde V.,
RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA Raoult D.,
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";

RL Science 293:2093-2098(2001).
DR EMBL: AE008609; AAL02990.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 312 AA; 34025 MW; A41878CAABCA9F5 CRC64;

Query Match 45.9%; Score 811.5; DB 16; Length 312;
Best Local Similarity 54.3%; Pred. No. 2.2e-44;
Matches 158; Conservative 60; Mismatches 72; Indels 1; Gaps 1;

OY 37 VLFVPOQEAAMVVERMGFRHILEPGLNLIIPVDRIRYVOSLKEIYINVPESAVTLDN 96
DB 19 MVRVFPQQAAMVVERKLEKFEKVLQPGNLILPIQIRVAYKHTLKEAIDVTAQTALISNDN 78
OY 97 VTLQIDGVLYLRIMDPYKASVGEDEPEYAVTOLAQTMRSELGKLSXDKVFRERESINAS 156
DB 79 VTLSDIGVLVKTIIDPAASVGVNPPYATVLAQTMRSELGKLPDRFEERETLNA 138
OY 157 IVDAINQADDCGICRLEIKDIHPPRVKESMOMQVAEKRRKATVLESEGTRESAIN 216
DB 139 IVAINQAINMGIOCMRYEIKDIQPPQITLKAMELQVAERQKRAQJLSEGNRAKIN 198
OY 217 VAEKKAQOILASEAEKAEQINQAGEASAVLAKKAKAEAIRILAAALTOHNGDAASL 276
DB 199 HAEGERKAQIVLNSEASVTDQVNRAGEAIGLVATFANSIEIVATVAYVKTGSDAVAL 258
OY 277 TVAEQYVSASFSLKARDSTNTLLPSNPGDVTSMVAQAMGYV-GALTKAPVPGT 327
DB 259 KIAEYISAFAGNLAKDNTVILPANLSEPSFTTEALTIFNOL-KASSSGT 308

RESULT 14
OY2DK0
ID O92DK0 PRELIMINARY; PRT; 311 AA.
AC O92DK0;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE HYPOTHETICAL 34.0 KDA PROTEIN.
GN RP328.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sichteritz-Ponten T., Almark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.,
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria.";
RL Nature 396:133-140(1998).
DR EMBL: AJ235271; CAI4788.1; -
DR InterPro: IPR001107; Band_7.
DR InterPro: IPR001972; Stomatlin.
DR Pfam: PF01145; Band_7; 1.
DR PRINTS: PR00721; STOMATIN.
DR SMART: SM00244; PHB; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 311 AA; 34040 MW; F74F625EC567E5CA CRC64;

Query Match 45.6%; Score 806; DB 16; Length 311;
Best Local Similarity 54.4%; Pred. No. 5e-44;
Matches 154; Conservative 60; Mismatches 69; Indels 0; Gaps 0;

OY 37 VLFVPOQEAAMVVERMGFRHILEPGLNLIIPVDRIRYVOSLKEIYINVPESAVTLDN 96
DB 19 MVRVFPQQAAMVVERKLEKFEKVLQPGNLILPIQIRVAYKHTLKEAIDVTAQTALISNDN 78
OY 97 VTLQIDGVLYLRIMDPYKASVGEDEPEYAVTOLAQTMRSELGKLSXDKVFRERESINAS 156

Db 79 VLLSLDGLVYKIIDPMASISYVNNPYAITQLAOTMSEIGKLPDTEFERDLNVA 138
 Oy 157 IVDAINQADDCGICRLREIDIHVPVRESQMOQVAERKRATVLESEGTRESAIN 216
 Db 139 IYSALINQAAINNGICQMRKEIKDIOPOTILKAMEIQVAEROKRAQIIESEGNQAKIN 198
 Oy 217 VAEGKKAQIILASEAKAQIINQAGEASAVLAKAKAKAIAIRILAAALTOHNGDAASL 276
 Db 199 HAEGEKAQIVLNSASYSYTDQVNRKAGEAIGLVATATANSIEIYMAAIQKTGSDAVAL 258
 Oy 277 TVAEQVSAFSKLAKDSNTILLPSNPGDVTSMVAQMGYGCAL 319
 Db 259 KIAEQYISAFGNLAKDVTNVLPTNLSEPSFVGTGALTIFNQL 301

Oy 351 DRVK 354
 Db 388 QRPK 391

Search completed: September 22, 2002, 18:30:03
 Job time: 546 sec

RESULT 15
 ID 09M585 PRELIMINARY: PRT: 394 AA.
 AC 09M585:
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
 DE 01-JUN-2001 (TREMBLrel. 17, last annotation update)
 DE STOMATIN-LIKE PROTEIN.
 GN STM1.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nadimpalli R., Yalpani N., Johal G.S., Simmons C.R.;
 RT "Prohibitins, Stomatins, and Plant Disease Response Genes Comprise a
 RT Protein Superfamily that Controls Cell Proliferation, Ion Channel
 RT Regulation, and Death."
 RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
 RL EMBL: AF236372; AAF68388.1; -.
 DR InterPro: IPR001107; Band_7.
 DR InterPro: IPR001972; Stomatlin.
 DR Pfam: PF01145; Band_7; 1.
 DR PRINTS: PR00721; STOMATIN.
 DR SMART: SM00244; PHB; 1
 SQ SEQUENCE 394 AA: 43277 MW: 52A409DCFI5C5A45 CRC64:

Query Match 45.0%; Score 794.5; DB 10; Length 394;
 Best Local Similarity 47.0%; Pred. No. 3.7e-43;
 Matches 171; Conservative 63; Mismatches 107; Indels 23; Gaps 3;

Oy 14 ABGLSTGFVPRSGRASSGL-----PRNTVLEFVPOQEAHVVERMGFRHRIIEPGNLILI 67
 Db 28 APALSRFSFRNPDDSSMFDPEPPVNMGSIVPEKKAIVVERFGKYLTGLSGFHLII 87
 Oy 68 PVLDRIRYVQSLEKIYINPEQSAVTLDNVTLQIDGVLYLRIMDPYKASGVEDPEYAVT 127
 Db 88 PAVDRIAVYHSLKEETIPPIPHONAIFKIDVNTIQIDSVIYVKIMDPYLASGVENPIYAVL 147
 Oy 128 QLAOTTMRSELGLKSDKVFRESINASTVDAINQAAOCMGIRCLRYETIKDTHVPPRVK 187
 Db 148 QLAOTTMRSELGKITTDKTFEERDALNEKIVSAINENATDMGLKCTRYETIRDINPPAGIR 207
 Oy 188 ESMQOVEAERKRATVLESEGTRESAINVAEGKKAQIILASEAKAQIINQAGEASAV 247
 Db 208 QAMEQQAEEARRKRAQIILESECKKQAOIIESECKTAQIIESEGAMLDLANRAKGAEEAI 267
 Oy 248 LAKAKAKAIAIRILAAALTOHNGDAASLTVAEQVSAFSKLAKDSNTILLPSNPGDVT 307
 Db 268 LAKSEATAGCMRLVSDAMTTEGSAKASLKLABQYIEAFSNLAQKNTMTMLPDGDSASPAS 327
 Oy 308 MVNAQMGVY-----GALTKAP-----VPGTPDLSGSSSRDVGCTDASXDEEL 350
 Db 328 FVAQAKKTYEQIHSHSQALSHPOIEBELKEGSETSPAPSSASKTPPLLEADSNQTFSL 387


```

RESULT      4
US-08-781-562-3
Sequence 3, Application US/08781562
Patent No. 5763589
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Goll, Surya K.
TITLE OF INVENTION: NOVEL HUMAN MEMBRANE PROTEIN
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,562
FILING DATE: Herewith
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0181 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 288 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 31069
? ? ?
US-08-781-562-3

```

```

1 GENERAL INFORMATION:
2 APPLICANT: Hillman, Jennifer L.
3 APPLICANT: Goll, Surya K.
4 TITLE OF INVENTION: NOVEL HUMAN MEMBRANE PROTEIN
5 NUMBER OF SEQUENCES: 7
6 CORRESPONDENCE ADDRESS:
7 ADDRESSEE: Incyte Pharmaceuticals, Inc.
8 STREET: 3174 Porter Drive
9 CITY: Palo Alto
10 STATE: CA
11 COUNTRY: USA
12 ZIP: 94304
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE: Diskette
15 COMPUTER: IBM Compatible
16 OPERATING SYSTEM: DOS
17 SOFTWARE: FASTSEQ for Windows Version 2.0
18 CURRENT APPLICATION DATA:
19 APPLICATION NUMBER: US/08/781,562
20 FILING DATE: Herewith
21 CLASSIFICATION: 530
22 PRIOR APPLICATION DATA:
23 APPLICATION NUMBER:
24 FILING DATE:
25 ATTORNEY/AGENT INFORMATION:
26 NAME: Billings, Lucy J.
27 REGISTRATION NUMBER: 36,749
28 REFERENCE/DOCKET NUMBER: PF-0181 US
29 TELECOMMUNICATION INFORMATION:
30 TELEPHONE: 415-855-0555
31 TELEFAX: 415-845-4166
32 TELEX:
33 INFORMATION FOR SEQ ID NO: 4:
34 SEQUENCE CHARACTERISTICS:
35 LENGTH: 280 amino acids
36 TYPE: amino acid
37 STRANDEDNESS: single
38 TOPOLOGY: linear
39 IMMEDIATE SOURCE:
40 LIBRARY: GenBank
41 CLONE: 1065452
42 US-08-781-562-4
43
44 Query Match 15.3%; Score 271; DB 1; Length 280;
45 Best Local Similarity 27.5%; Pred. No.1.le-18;
46 Matches 73; Conservative 64; Mismatches 76; Indels 50; Gaps 8
47
48 41 VQGGAWVERNGRF--HRLLEPGNLILIPVLDRIYVQSLKEIVINPEQSAVTLDNVT 98
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

```


	Matches	356;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	MLARAARGHWGPFACGLSTGFWPRSGRASSGGLPRNTVYLFVWQOEAWVVEFGREHRIIE	60							
Db	1	MLARAARGHWGPFACGLSTGFWPRSGRASSGGLPRNTVYLFVWQOEAWVVEFGREHRIIE	60							
QY	61	PLGLNTLIVLDRIRVQSKETIVNWPEQSAATLNLNTLQIDGVLVYLRIMDPKASGYE	120							
Db	61	PLGLNTLIVLDRIRVQSKETIVNWPEQSAATLNLNTLQIDGVLVYLRIMDPKASGYE	120							
QY	121	DPEYAVTOLAQTMRSELGLKLSXDRVFERRESLNASIVDAINQADDCWGIICFLRYEIKDI	180							
Db	121	DPEYAVTOLAQTMRSELGLKLSXDRVFERRESLNASIVDAINQADDCWGIICFLRYEIKDI	180							
QY	181	HVPFVKESMOMQVBAERRKRATVLESECTRRESAINVAEGKKQAQIILASEAKAEQINQA	240							
Db	181	HVPFVKESMOMQVBAERRKRATVLESECTRRESAINVAEGKKQAQIILASEAKAEQINQA	240							
QY	241	AGEASAVLAKKAKAEAIRIILAAALTQHNHGDAAASLTVAEQVYSAFSKIAKDSNTILLES	300							
Db	241	AGEASAVLAKKAKAEAIRIILAAALTQHNHGDAAASLTVAEQVYSAFSKIAKDSNTILLES	300							
QY	301	NPGDVTSWVAQMGVYGAITKAPVGTGPTSLSSGSSRDVGQTDASDEEDLDRYKMS	356							
Db	301	NPGDVTSWVAQMGVYGAITKAPVGTGPTSLSSGSSRDVGQTDASDEEDLDRYKMS	356							

[illegible]

```

1      RESULT 2
2      US-08-781-562-6
3      : Sequence 6, Application US/08781562
4      : Patent No. 5763589
5      :
6      : GENERAL INFORMATION:
7      : APPLICANT: Hillman, Jennifer L.
8      : APPLICANT: Goll, Surya K.
9      : TITLE OF INVENTION: NOVEL HUMAN MEMBRANE PROTEIN
10     : NUMBER OF SEQUENCES: 7
11     : CORRESPONDENCE ADDRESS:
12     : ADDRESSEE: Incyte Pharmaceuticals, Inc.
13     : STREET: 3174 Porter Drive
14     : CITY: Palo Alto
15     : STATE: CA
16     : COUNTRY: USA
17     : ZIP: 94304
18     :
19     : COMPUTER READABLE FORM:
20     : MEDIUM TYPE: Diskette
21     : COMPUTER: IBM Compatible
22     : OPERATING SYSTEM: DOS
23     :
24     : SOFTWARE: FastSeq for Windows Version 2.0
25     :
26     : CURRENT APPLICATION DATA:
27     : APPLICATION NUMBER: US/08/781,562
28     : FILING DATE: Herewith
29     : CLASSIFICATION: 530
30     : PRIOR APPLICATION DATA:
31     : APPLICATION NUMBER:
32     : FILING DATE:
33     :
34     : ATTORNEY/AGENT INFORMATION:
35     : NAME: Billings, Lucy J.
36     : REGISTRATION NUMBER: 36,749
37     : REFERENCE/DOCKET NUMBER: PF-0181 US
38     : TELECOMMUNICATION INFORMATION:
39     : TELEPHONE: 415-855-0555
40     : TELEFAX: 415-845-4166
41     :
42     : TELEX:
43     :
44     : INFORMATION FOR SEQ ID NO: 6:
45     : SEQUENCE CHARACTERISTICS:
46     : LENGTH: 381 amino acids
47     : TYPE: amino acid
48     : STRANDEDNESS: single
49     : TOPOLOGY: linear
50     : IMMEDIATE SOURCE:
51     : LIBRARY: OMI
52     : CLONE: 79701
53     :
54     : US-08-781-562-6

```

RESULT 3
 US-08-781-562-7
 Sequence 7, Application US/08781562
 Patent No. 5763589
 GENERAL INFORMATION:
 APPLICANT: Hillman, Jennifer L.
 APPLICANT: Goli, Suyra K.
 TITLE OF INVENTION: NOVEL HUMAN MEMBRANE PROTEIN
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/781,562
 FILING DATE: Herewith
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0181 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 TELEX:
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 199 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 17:24:54 : Search time 24.86 Seconds
(without alignments)
349.779 Million cell updates/sec

Title: US-09-898-216-1

Perfect score: 1767
Sequence: 1 MLRAARGHMGFAEGSLTG.....RDVGTDSXDELDYRWMS 356

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/ECTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1763	99.8	356	1	US-08-781-562-1
2	472	26.7	381	1	US-08-781-562-6
3	354.5	20.1	199	1	US-08-781-562-7
4	281.5	15.9	288	1	US-08-781-562-3
5	271	15.3	280	1	US-08-781-562-4
6	169	9.6	415	1	US-08-781-562-5
7	154	8.7	379	1	US-08-279-270A-1
8	104.5	5.9	516	2	US-08-762-106-8
9	104.5	5.9	516	3	US-08-745-404-2
10	104.5	5.9	516	4	US-09-320-774-8
11	104.5	5.9	527	2	US-08-762-106-9
12	104.5	5.9	527	4	US-09-320-774-9
13	104.5	5.9	552	3	US-08-745-404-3
14	104	5.9	1339	4	US-09-310-187A-1
15	99.5	5.6	564	2	US-08-115-746-2
16	99.5	5.6	564	4	US-09-115-746-2
17	99.5	5.6	643	2	US-08-216-894-8
18	99.5	5.6	643	4	US-09-115-746-8
19	99.5	5.6	1886	4	US-08-938-105-3
20	99	5.6	1792	2	US-08-962-284-4
21	97.5	5.5	472	2	US-08-216-894-10
22	97.5	5.5	472	4	US-09-115-746-10
23	97.5	5.5	1196	1	US-08-144-121-4
24	97.5	5.5	1196	2	US-08-735-893-4
25	96.5	5.5	400	3	US-08-938-830-29
26	96.5	5.5	914	4	US-09-085-199B-4
27	95.5	5.4	3170	2	US-07-642-734C-5

28	95.5	5.4	3170	3	US-08-439-009A-5	Sequence 5, Appl1
29	95	5.4	1184	4	US-09-541-782-2	Sequence 2, Appl1
30	93.5	5.3	656	1	US-08-371-930-23	Sequence 23, Appl1
31	93.5	5.3	656	5	PCT-US84-01712-23	Sequence 23, Appl1
32	93.5	5.3	676	1	US-08-371-930-24	Sequence 24, Appl1
33	93.5	5.3	676	5	PCT-US94-01712-24	Sequence 24, Appl1
34	93.5	5.3	694	1	US-08-339-152A-18	Sequence 18, Appl1
35	93.5	5.3	694	2	US-08-007-999B-5	Sequence 5, Appl1
36	93.5	5.3	694	2	US-08-689-276A-5	Sequence 5, Appl1
37	93.5	5.3	695	1	US-08-371-930-27	Sequence 27, Appl1
38	93.5	5.3	695	1	US-08-123-702-2	Sequence 2, Appl1
39	93.5	5.3	695	1	US-08-339-152A-30	Sequence 30, Appl1
40	93.5	5.3	695	2	US-08-104-165-1	Sequence 1, Appl1
41	93.5	5.3	695	3	US-08-464-250-1	Sequence 1, Appl1
42	93.5	5.3	695	4	US-08-464-250-1	Sequence 1, Appl1
43	93.5	5.3	695	4	US-09-458-481B-7	Sequence 7, Appl1
44	93.5	5.3	695	4	US-09-458-481B-8	Sequence 8, Appl1
45	93.5	5.3	695	5	PCT-US94-01712-27	Sequence 27, Appl1

ALIGNMENTS

RESULT 1
US-08-781-562-1
; Sequence 1, Application US/08781562
; Patent No. 5763589
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goll, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN MEMBRANE PROTEIN
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,562
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0181 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 356 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: Consensus
; US-08-781-562-1
Query Match 99.8%; Score 1763; DB 1; Length 356;
Best Local Similarity 100.0%; Pred. No. 2.6e-165;

Db 45 sagypnsfqltpptngiivperkafvierfgkyaltlpsgihlfpvdrtayvhs1 104
Qy 80 KEIYINPEOSAVYLDNVTQIDSVLYLRIMDPYKASYGVDEPEYAVTOLAQTMRSELG 139
Db 105 keeaiipnqtatkhvshidgvlvkvikvdplasygvespiyavqiaqltmrse1g 164
Qy 140 KLSXKVFRRERESINASYDAINQADCWGIRCLRYEIKIDIHPPRVKESMQOVEERR 199
Db 165 kltldkfeerdltnekeiveaveakdwglqclryelrldmphyvraamemqaeerk 224
Qy 200 KRAATVESEGTREATAINWAGKKKAOQLASFAEKAEQINQAGASAVLAKAKAKAEAIR 259
Db 225 kraqliesegershniadgkssvllaaseakmdqvnraqaeaalaraqatakglv 284
Qy 260 ILAALTOHNGDAASLTVAEQYVSASFSLAKDSNTILPSPGDTSMVAQAMGVYCAL 319
Db 285 llsgslteygvaeaslrveeqyitafgniakegtlmlpsgasnpsamlaqaltmks1 344
Qy 320 TKAFVPGTSPSLSSGSSRDVQGTDAKXDELDKRVK 354
Db 345 -----vingspkdhqetqaldetdele 368

RESULT 15
AAG48817
ID AAG48817 standard; Protein: 382 AA.
XX AAG48817;
AC
XX
XX 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 61686.
XX
KM Protein identification: signal transduction pathway; metabolic pathway;
KM hybridisation assay: genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.


```
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 46.7%; Score 826; DB 21; Length 411;

Best Local Similarity 50.1%; Pred. No. 4, 5e-64;

Matches 168; Conservative 67; Mismatches 80; Indels 20; Gaps 2;

```
OY 29 SSGLPNTVVL-----FVPOQEAWYVERGRHRIIEPRLNLIPLDLRIRVQSL 79
DB 45 saqyspsnfqlprrptmgwlrirprrkafvrrkyatclprrghlrlprrvdlrlyvhsll 104
OY 80 KEIVIVPQSAVTLDDNVTLOIDGVLYLRIMDPYKASVGEVDEPEVAVTOLAQJTTMSELC 139
DB 105 keeaiprpqtaiktksvnhidgvlrvkltvdpkrlasgyvespiyavvglaqtlmrselg 164
OY 140 KLSXKQVFRRESLNASIYDAINQAADCGIRCLREIKDIIHVPYKESMNOVEAERR 199
DB 165 kltdktlfeerdlinekivealnvaaakdwlglcrlryelrldimphgvraamemqaeaeerk 224
OY 200 KRATVESEGTRESATINVAEGKQAOIILASEAKAOINQAGEASAVLAKAKAEAIR 259
DB 225 kraqllesegrrqshnladgkssvllaseaakmdqvmragaealllaraqatakglv 284
```

```
OY 260 ILAALTOHNGDAASLTVAEOVYSAFSLAKDSNTILLPSNPGDVTSMVAQMGVYCAL 319
DB 285 llsqslketgyveaaslrvaeyrtalfgnlaketclmlpssgasnpsmalaqaltmlyksl 344
OY 320 TKAPVGPDPDSLSSGSSRPDVCGTDAKXDELDPRVK 354
DB 345 -----vngpskdhgetqaldecldleele 368

RESULT 14
AAG31395
ID AAG31395 standard; Protein: 515 AA.
XX
AC AAG31395;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 37695.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-01231825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130409.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136592.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
```

XX AAG50184;
AC 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 63567.
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
OS Arabidopsis thaliana.
XX EPI033405-A2.
XX 06-SEP-2000.
PD 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.

CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
CC activity, regulation of hematopoiesis and is useful for treating myeloid
CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
CC tissue growth, and in tissue repair, healing of burns, incisions,
CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
CC disorders, or periodontal disease. Furthermore, (I) is also useful for
CC gut protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, various immune deficiencies and
CC disorders including severe combined immunodeficiency (SCID), bacterial or
CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
CC reactions and conditions, such as asthma or other respiratory problems.
CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
CC fertility, metabolism, catabolism, anabolism, storage or elimination of
CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
CC analgesic effects or other pain reducing effects, immunoglobulin like
CC activity and can act as an antigen in a vaccine composition to raise an
CC immune response. AAU28020-AAU28395 represent novel human secreted protein
CC amino acid sequences of the invention.

XX Sequence 463 AA:

Query Match 53.5%; Score 946; DB 22; Length 463;
Best Local Similarity 64.0%; Pred. No. 1,6e-74;
Matches 242; Conservative 20; Mismatches 74; Indels 42; Gaps 19;

OY 13 FAEG--LSTGFMFRSGRASSGLPRNTVLFVPOEAMVVRMRGPHRLPGLNIIIPVL 70
DB 94 fLggsllaaagrapr--rassgdlprncvllfpqgaawvermrtrfhllepjnlhlpvl 151
OY 71 DRIRYV-QSLKETIV-INVP-EGSAYLTDNVTLQIDGVLYLRINDPYKAS-YGEDEPEYAV 126
DB 152 drrlyaaesskeiclsfnvalnsavcldnvclqldgvlylrindpykqkrlgyedpeyav 211
OY 127 TQLAQT-MRSELGKLSXKXVFERESLNASYDAINQAAOCGICRLREIDIHPPR 165
DB 212 lqpslnhvdqsglqfslqfreslnaslydaIngaacwglrclyelkdhppr 271
OY 186 VKESMOMOVEAERR-----KRAVLESESG--TRESAINVAGCKKQAOVLASEAEKAEQ 236
DB 272 vke-----vxagmhggggqgqgktratlvesegdpcksaqfwgeqkqgqlaseaeaeq 326
OY 237 INQAGEASAVLAKAR----AKAEAIRLAAALTOHNGDPAASLT--VAE--QYVSASF 287
DB 327 insgqerascsgerprkllkairfgrlqaldpknwelaafxlvaeaamfsgvfp 386
OY 288 KLA--KDSMTI--LLEPSNEDVTSMAQAMGV-YGALTKAPVPGTDP---SLSSGSRDY 339
DB 387 klqprnsgqlplclpct-pqdvtsmwaaqpwvxygaltspspardfqlhssssardv 445
OY 340 OGTDAS-XDELDPRVKMS 356
DB 446 ggtdeaspxxgnllrvkms 463

RESULT 12

AAG20661

ID AAG20661 standard: Protein; 411 AA.

AC AAG20661;

XX 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 22939.

KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 9905-0121825.

PR 05-MAR-1999; 9905-0123180.

PR 09-MAR-1999; 9905-0123548.

PR 23-MAR-1999; 9905-0125788.

PR 25-MAR-1999; 9905-0126264.

PR 29-MAR-1999; 9905-0126785.

PR 01-APR-1999; 9905-0127462.

PR 06-APR-1999; 9905-0128234.

PR 08-APR-1999; 9905-0128714.

PR 16-APR-1999; 9905-0129845.

PR 19-APR-1999; 9905-0130077.

PR 21-APR-1999; 9905-0130449.

PR 23-APR-1999; 9905-0130510.

PR 23-APR-1999; 9905-0130891.

PR 28-APR-1999; 9905-0131449.

PR 30-APR-1999; 9905-0132048.

PR 30-APR-1999; 9905-0132407.

PR 04-MAY-1999; 9905-0132484.

PR 05-MAY-1999; 9905-0132485.

PR 06-MAY-1999; 9905-0132486.

PR 07-MAY-1999; 9905-0132487.

PR 11-MAY-1999; 9905-0132863.

PR 14-MAY-1999; 9905-0134256.

PR 14-MAY-1999; 9905-0134218.

PR 14-MAY-1999; 9905-0134219.

PR 14-MAY-1999; 9905-0134221.

PR 18-MAY-1999; 9905-0134370.

PR 19-MAY-1999; 9905-0134768.

PR 19-MAY-1999; 9905-0134841.

PR 20-MAY-1999; 9905-0135124.

PR 21-MAY-1999; 9905-0135353.

PR 24-MAY-1999; 9905-0135629.

PR 25-MAY-1999; 9905-0136021.

PR 27-MAY-1999; 9905-0136392.

PR 28-MAY-1999; 9905-0136782.

PR 01-JUN-1999; 9905-0137222.

PR 03-JUN-1999; 9905-0137528.

PR 04-JUN-1999; 9905-0137528.

PR 07-JUN-1999; 9905-0137724.

PR 08-JUN-1999; 9905-0138094.

PR 10-JUN-1999; 9905-0138540.

PR 10-JUN-1999; 9905-0138847.

PR 14-JUN-1999; 9905-0139119.

PR 16-JUN-1999; 9905-0139452.

PR 16-JUN-1999; 9905-0139453.

PR 17-JUN-1999; 9905-0139454.

PR 18-JUN-1999; 9905-0139452.

PR 18-JUN-1999; 9905-0139455.

PR 18-JUN-1999; 9905-0139456.

PR 18-JUN-1999; 9905-0139457.

PR 18-JUN-1999; 9905-0139458.

PR 18-JUN-1999; 9905-0139459.

PR 18-JUN-1999; 9905-0139460.

PR 18-JUN-1999; 9905-0139461.

PR 18-JUN-1999; 9905-0139462.

PR 18-JUN-1999; 9905-0139463.

PR 18-JUN-1999; 9905-0139750.

PR 21-JUN-1999; 9905-0139763.

PR 22-JUN-1999; 9905-0139817.

PR 23-JUN-1999; 9905-0139899.

PR 23-JUN-1999; 9905-0140353.

PR 24-JUN-1999; 9905-0140354.

PR 24-JUN-1999; 9905-0140695.

PR 28-JUN-1999; 9905-0140825.

PR 29-JUN-1999; 9905-0140981.

PR 30-JUN-1999; 9905-0141287.

PR 01-JUL-1999; 9905-0141842.

OY 339 VQGTDA-----SXDEELDRVKMS 356
 Db 294 vkvgagclnaksveykkelgedksvkmn 321

RESULT 10
 AAU33158
 ID AAU33158 standard; Protein: 222 AA.

AC AAU33158:

DT 18-DEC-2001 (first entry)

DE Novel human secreted protein #3649.

KW Human: vaccination; gene therapy; nutritional supplement;
 stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 immune suppression; immune stimulation; anti-inflammatory; leukaemia.

OS Homo sapiens.

PN WO200179449-A2.

PD 25-OCT-2001.

PF 16-APR-2001; 2001WO-US08656.

PR 18-APR-2000; 2000US-0552929.

PR 26-JAN-2001; 2001US-0770160.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI: 2001-611725/70.

PT Nucleic acids encoding a range of human polypeptides, useful in genetic
 vaccination, testing and therapy -

PS Claim 20; Page 716; 765pp; English.

CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.

XX Sequence 222 AA:

Query Match 57.2%; Score 1011; DB 22; Length 222;

Best local similarity 92.9%; Pred. No. 1, 1e-80;

Matches 208; Conservative 2; Mismatches 10; Indels 4; Gaps 2;

OY 1 MLARAARGHMCPFAEG--LSTGFMPRSGRASSGIPRNTVVLFFVPOOEAVVBERMGRFHRI 58

Db 1 mlaraargtgallrsgllasgrapr--rassglprntvtvlffvpggeavwvermgtrhri 58

OY 59 LEPGLNLLPVDLRIRYVOSLKEIVINVPQSAVTLDNVTLOIDGVLYLRIMDPYKASY 118

Db 1 lepglnllpvdldirryvsgslkeivinvpqsavtldnvtldidgvllylrimdpykasy 118

Db 59 lepglnllpvdldirryvsgslkeivinvpqsavtldnvtldidgvllylrimdpykasy 118
 OY 119 VEDPEYAVTOLAQTMRSELKLSXDKVFERERESINASTVADINAOADQWGRICRLRYEIK 178
 Db 119 vedpeyavtqlaqtlmrseqlklsdkvferesinasivadinadqowgricrlryeik 178
 OY 179 DIHPPRYKESHQMOQVEAEERRKATVLESEGTRESAINVAEKK 222
 Db 179 dihprrykesmqmqveaerkratvlesegtresainvaeqk 222

RESULT 11

AAU28241
 ID AAU28241 standard; Protein: 463 AA.

AC AAU28241:

DT 18-DEC-2001 (first entry)

DE Novel human secretory protein, Seq ID No 598.

KW Human: secreted protein; arthritis; Crohn's disease; sepsis; shock;
 KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KW fertility; analgesic; pain; antigen.

OS Homo sapiens.

PN WO200166689-A2.

PD 13-SEP-2001.

PF 05-MAR-2001; 2001WO-US04942.

PR 07-MAR-2000; 2000US-0519705.

PR 19-MAY-2000; 2000US-0574454.

PR 17-JUN-2000; 2000US-0596193.

PR 14-JUL-2000; 2000US-0618447.

PR 19-SEP-2000; 2000US-0665363.

PR 20-OCT-2000; 2000US-0693267.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
 PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;

DR WPI: 2001-589934/66.

PT N-PSDB: AAS45141.

PT Novel polypeptides and nucleic acids obtained from cDNA libraries
 prepared from various human tissues, for diagnosis and treatment of
 cancer, neurological, inflammatory, and autoimmune disorders -

PS Example 2; SEQ ID No 598; 107pp; English.

CC The invention relates to novel isolated human secreted polypeptides (I)
 CC and polynucleotides (II). (I) and (II) are useful for treating
 CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
 CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
 CC involved in increasing haematopoiesis, stem cell survival, bone growth
 CC and remodeling. (I), (II) and modulators of (II) are useful for
 CC creating transgenic animals useful for studying the in vivo activities of
 CC the polypeptide as well as for studying modulators of the polypeptides.
 CC (I) induces the proliferation of neural cells and regeneration of nerve
 CC and brain tissue and is useful for the treatment of central and
 CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
 CC Parkinson's disease, Huntington's disease, and amyotrophic lateral

Qy 299 PSNPGDVTSMVAQAMGYGALTKAPVPGTSPDLSISGSSRDVGTDPASXDEELDRVKMS 356
|||||
Db 299 psnpgdvtsmwagamygyaltkapyvpgtspdlsissgsrdvgtgtdasldaeeldrvkmt 356

RESULT 7

AAW78566
ID AAW78566 standard; Protein: 358 AA.

AC AAW78566;

DT 06-NOV-2001 (first entry)

DE Human protein SEQ ID NO 1228.

XX Human: cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.

OS Homo sapiens.

PN MO200157190-A2.

PD 09-AUG-2001.

PF 05-FEB-2001; 2001WO-US04098.

PR 03-FEB-2000; 2000US-0496914.

PR 27-APR-2000; 2000US-0560875.

PR 20-JUN-2000; 2000US-0598075.

PR 19-JUL-2000; 2000US-0620325.

PR 01-SEP-2000; 2000US-0654936.

PR 15-SEP-2000; 2000US-0663561.

PR 20-OCT-2000; 2000US-0693325.

PR 30-NOV-2000; 2000US-0728422.

DR N-PSDB; AAK51699.

XX WPI: 2001-476283/51.

XX N-PSDB; AAK51699.

XX Nucleic acids encoding polypeptides with cytokine-like activities;

XX useful in diagnosis and gene therapy -

XX Claim 20; Page 3486-3487; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAW78323-AAW80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAW80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.

XX Sequence 358 AA;

Query Match 92.5%; Score 1635; DB 22; Length 358;
Best Local Similarity 94.7%; Pred. No. 4,1e-135;
Matches 341; Conservative 2; Mismatches 11; Indels 6; Gaps 3;

Qy 1 MLARAAGHMGPFACG--LSTGFMPSRSGASSGLPRTNVTVLEVPQOEAAMVVERMGFRHRI 58
|||||
Db 1 mlaraargtgalllrgslsaaigrpr--raasgldprntvlfvpgqaeawvermgfrfni 58

Qy 59 LEPGLNLIPLVDRIYVQSLKEIVINVEQSAVTLNDVTLQIDGVLYRLIMDPKASYG. 118
|||||

Db 59 lepgnlllpvldrlirgyslkeivinvpeqsavldnvtlqldgvlyrlimdpkasyg 118

Qy 119 VEDPEYATQLAQTMRSELGKLSXDKVFRFRESINASTIVAINAACCWGRICLRFEIK 178
|||||

Db 119 vedpeyavtlqqlqtmrseqlklsldkvfrereslnaastivadnqadawgriclrlyek 178

Qy 179 DIHVPPRKESMOMQVEARRKATVLESEGTRESAIVNAEGRKQAOILASEAEKAEQIN 238
|||||

Db 179 dihvpprvkesmqvvearrkratvlesegtrresainvaeqkkaqqlaaseakeaqin 238

Qy 239 QAGGASAVLAKAKAKAEIRILAAALROHNGDAASLTVAEQVYSAPSKLAKDSNTILL 298
|||||

Db 239 qaagsasavlakakakaairllaaaltqhngdaaastlvaeeqysatsklakdsntlll 298

Qy 299 PSNPGDVTSMVAQAMGYGALTKA--PYGTPDLSISGSSRDVGTDPASXDEELDRVKMS 356
|||||

Db 299 psnpgdvtsmwagamygyaltkapyvpgtspdlsissgsrdvgtgtdasldaeeldrvkms 356

RESULT 8

AAW78284
ID AAW78284 standard; Protein: 306 AA.

AC AAW78284;

DT 13-APR-1999 (first entry)

DE Fragment of human secreted protein encoded by gene 35.

XX Human: secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

OS Homo sapiens.

PN WO9856804-A1.

PD 17-DEC-1998.

PF 11-JUN-1998; 98WO-US12125.

PR 02-OCT-1997; 97US-0061060.

PR 13-JUN-1997; 97US-0049547.

PR 13-JUN-1997; 97US-0049548.

PR 13-JUN-1997; 97US-0049549.

PR 13-JUN-1997; 97US-0049550.

PR 13-JUN-1997; 97US-0049606.

PR 13-JUN-1997; 97US-0049607.

PR 13-JUN-1997; 97US-0049608.

PR 13-JUN-1997; 97US-0049609.

PR 13-JUN-1997; 97US-0049610.

PR 13-JUN-1997; 97US-0050566.

PR 13-JUN-1997; 97US-0050901.

PR 08-JUL-1997; 97US-0052989.

PR 18-AUG-1997; 97US-0051919.

PR 12-SEP-1997; 97US-0058665.

PR 12-SEP-1997; 97US-0058666.

PR 12-SEP-1997; 97US-0058668.

PR 12-SEP-1997; 97US-0058669.

PR 12-SEP-1997; 97US-0058750.

Db 1 mlaraatgctgalllrgslasgrpr--raasgiprntvlfvpqgeawvermgrrftrci 58
 QY 59 LEPLNTLIPVLRIRYVOSLKEIVINWPEQSAVTLDNWTLQIDGVLYLRIMDPYKASYG 118
 Db 59 lrpqnlilipvldrlryvgsikeivinvpegsavtlndvldqgvllylrmdpykasyg 118
 QY 119 VEDPEYAVTQLAOTMTSELGKLSXDKVFRRESLNMSYDAIQOAAADCMGICRLREIK 178
 Db 119 vedpeyavtqlagctlmseigklsldkfrereslnasivdaingaadcwglrclryeik 178
 QY 179 DIHVPKESMOMQVFAERRRRATVLESEGTRESALINVEGKQAOILSEAKAKOIN 238
 Db 179 dihvprvkemqmvaeerrkratvlesegtresalnveegkqagqlaseakeeqin 238
 QY 239 QAAGEASAVLAKAKAKAFAIRIILAAALTOHNGDAASLTVAEOYVSAFSKLANDSNTLL 298
 Db 239 qaageasaavilakakakeaairilaaaltqngdaaasltvaeqyvsafrsklaksntlll 298
 QY 299 PSNPGDVTSMVAQAMGYGALTRKAPVPGTDPDSLSSGSSRDYQGTDAEXDELDVKKMS 356
 Db 299 psnpgdvtsmvadamgyvgaltrkapygtpdsslssgsrdyvgtdaexldeeldrvkms 356

RESULT 3
 AAB92659
 ID AAB92659 standard; Protein: 356 AA.
 AC AAB92659;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:11010.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 OS Homo sapiens.
 XX
 PN EPI074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI: 2001-318749/34.
 XX
 PT primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 CC Claim 8: SEQ ID 11010; 2537pp + CD ROM; English.
 CC
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 356 AA;
 XX

Query Match 93.2%; Score 1646; DB 22; Length 356;
 Best Local Similarity 95.3%; Pred. No. 4,4e-136;
 Matches 341; Conservative 2; Mismatches 11; Indels 4; Gaps 2;

QY 1 MLARARGHMGCPFAEG--LSTGFMPRSGRASGSPRTVTVLFVPOQAMVVERMGRRFRI 58
 Db 1 mlaraargctgalllrgslasgrpr--raasgiprntvlfvpqgeawvermgrrftrci 58
 QY 59 LEPLNTLIPVLRIRYVOSLKEIVINWPEQSAVTLDNWTLQIDGVLYLRIMDPYKASYG 118
 Db 59 lrpqnlilipvldrlryvgsikeivinvpegsavtlndvldqgvllylrmdpykasyg 118
 QY 119 VEDPEYAVTQLAOTMTSELGKLSXDKVFRRESLNMSYDAIQOAAADCMGICRLREIK 178
 Db 119 vedpeyavtqlagctlmseigklsldkfrereslnasivdaingaadcwglrclryeik 178
 QY 179 DIHVPKESMOMQVFAERRRRATVLESEGTRESALINVEGKQAOILSEAKAKOIN 238
 Db 179 dihvprvkemqmvaeerrkratvlesegtresalnveegkqagqlaseakeeqin 238
 QY 239 QAAGEASAVLAKAKAKAFAIRIILAAALTOHNGDAASLTVAEOYVSAFSKLANDSNTLL 298
 Db 239 qaageasaavilakakakeaairilaaaltqngdaaasltvaeqyvsafrsklaksntlll 298
 QY 299 PSNPGDVTSMVAQAMGYGALTRKAPVPGTDPDSLSSGSSRDYQGTDAEXDELDVKKMS 356
 Db 299 psnpgdvtsmvadamgyvgaltrkapygtpdsslssgsrdyvgtdaexldeeldrvkms 356

RESULT 4
 AAB11981
 ID AAB11981 standard; peptide: 378 AA.
 AC AAB11981;
 XX
 DT 11-JAN-2002 (first entry)
 XX
 DE Human SLP-2 homologue, SEQ ID NO:2351.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolytic; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antisthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulneryary; antilucer.
 OS Homo sapiens.
 XX
 PN WO200157188-A2.
 XX

XX G011 SK, Hillman JL;
XX MPI: 1998-347418/30.
DR N-PSDB: AAV28867.
XX
XX DNA encoding human integral membrane protein - useful for producing
XX recombinant protein, for treatment of anaemia and cancer
XX
XX Claim 1: Column 35-38; 33pp; English.
XX
XX The present sequence represents human integral membrane protein (IMP).
XX IMP may be administered to a subject to treat disorders associated
XX with abnormal ion transport or membrane conductance as well as a
XX variety of tumours, e.g. haemolytic anaemias and prostatic, breast and
XX pancreatic tumours. A vector capable of expressing IMP, or a fragment
XX or a derivative thereof, may also be administered to a subject to treat
XX the haemolytic anaemias and prostatic, breast and pancreatic tumours.
XX
XX Sequence 356 AA:
SQ
Query Match 99.8%; Score 1763; DB 19; Length 356;
Best Local Similarity 100.0%; Pred. No. 2,4e-146;
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLARAAGHMGFPAEG-ISTGFMPRSGRASSGLPRNTVVLFPQOEAAMVERMGFRHRI 60
DB 1 mlaraaghwgpfagelstgftprsgassglprntvvlfpqgeawvermgfrhrlle 60
QY 61 PGLNIIIPVLDRIKRYVQSLKEIVINVEQSAVTLQIDGVLVLRIMDPKASGYE 120
DB 61 pgliniiplvldrikyvqslkeivinvpegsavtldnvlqldgvlvlylrmdpykasyge 120
QY 121 DEEYANTOLAOITMRESEGLKSDKVFRESEINASTIVAINOAAACGRCIRYEIKKI 180
DB 121 deeyantqlagltmrseglksdkvfreresinasivaingaadcgwrlrclyeikci 180
QY 181 HPPPRKESMOMQVEARRKRAVLESECTRESAINVABEGKQAOULASEAEKAEQINO 240
DB 181 hppprkesmqmvearrkratvlesectresainvabegkkaqlaseaekeaqinqa 240
QY 241 AGEASAVLAKAKAKAPATITLAAALTOHNGDAASLTVEQVYSASFSLAKISNTILPS 300
DB 241 ageasavlakakakeaaitlaaalqhngdaaasltveeqysatsklaknsntilps 300
QY 301 NPGDVSMVAQAMGYGALTAKAPVPGTPSLSSGSRDVGOTDASXDELDKVKMS 356
DB 301 npgdvsmvaqamgygaltakapvpctpslssgsrtdvgldasxdeeldivkms 356
RESULT 2
AAW78160
ID AAW78160 standard; Protein: 356 AA.
AC AAW78160;
XX
XX 13-APR-1999 (first entry)
DE Human secreted protein encoded by gene 35 clone HTXCS21.
XX
XX Human: secreted protein; fusion protein; gene therapy; protein therapy;
XX diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
XX developmental abnormality; foetal deficiency; blood; allergy; renal;
XX immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
XX inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
XX cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
XX osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
XX endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
XX Homo sapiens.
XX
XX MO9856804-AI.
PN

XX 17-DEC-1998.
PD
XX 11-JUN-1998; 98WO-US12125.
PF
XX 02-OCT-1997; 97US-0061060.
PR 13-JUN-1997; 97US-0049547.
PR 13-JUN-1997; 97US-0049548.
PR 13-JUN-1997; 97US-0049549.
PR 13-JUN-1997; 97US-0049550.
PR 13-JUN-1997; 97US-0049606.
PR 13-JUN-1997; 97US-0049607.
PR 13-JUN-1997; 97US-0049608.
PR 13-JUN-1997; 97US-0049609.
PR 13-JUN-1997; 97US-0049610.
PR 13-JUN-1997; 97US-0049611.
PR 13-JUN-1997; 97US-0050566.
PR 13-JUN-1997; 97US-0050901.
PR 13-JUN-1997; 97US-0052989.
PR 08-JUL-1997; 97US-0051919.
PR 18-AUG-1997; 97US-0055984.
PR 12-SEP-1997; 97US-0058665.
PR 12-SEP-1997; 97US-0058668.
PR 12-SEP-1997; 97US-0058669.
PR 12-SEP-1997; 97US-0058750.
PR 12-SEP-1997; 97US-0058971.
PR 12-SEP-1997; 97US-0058972.
PR 12-SEP-1997; 97US-0058975.
PR 02-OCT-1997; 97US-0060834.
PR 02-OCT-1997; 97US-0060841.
PR 02-OCT-1997; 97US-0060844.
PR 02-OCT-1997; 97US-0060855.
PR 02-OCT-1997; 97US-0061059.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Brewer LA, Ehnert R, Ferris AM, Feng P, Greene JM, Lafleur DM;
PI Moore PA, NI J, Olsen HS, Rosen CA, Ruben SM, Shi Y, Young P;
PI Yu GL;
XX
XX MPI: 1999-080881/07.
DR N-PSDB: AAX04345.
XX
XX New isolated human genes and the secreted polypeptides they encode -
XX useful for diagnosis and treatment of e.g. cancers, neurological
XX disorders, immune diseases, inflammation or blood disorders
XX
XX Claim 11: Page 282-283; 380pp; English.
XX
XX This sequence represents a secreted human protein encoded by the gene
XX clone detailed in the descriptor line. The gene can be used to generate
XX fusion proteins by linking to the gene to a human immunoglobulin Fc
XX portion (e.g. AAX04302) for increasing the stability of the fused
XX protein as compared to the human protein only.
XX
XX The invention relates to 86 novel genes and their fragments (nucleic
XX acid sequences: AAX04311-X04410; amino acid sequences AAW78126-W78225)
XX which are useful for preventing, treating or ameliorating medical
XX conditions e.g. by protein or gene therapy. Also, pathological
XX conditions can be diagnosed by determining the amount of the new
XX polypeptides in a sample or by determining the presence of mutations in
XX the new polynucleotides. Specific uses are described for each of the 86
XX polynucleotides, based on which tissues they are most highly expressed in
XX (see AAX04311 for described uses).
XX
XX Sequence 356 AA:
SQ
Query Match 93.2%; Score 1646; DB 20; Length 356;
Best Local Similarity 93.3%; Pred. No. 4,4e-136;
Matches 341; Conservative 2; Mismatches 11; Indels 4; Gaps 2;
QY 1 MLARAAGHMGFPAEG-ISTGFMPRSGRASSGLPRNTVVLFPQOEAAMVERMGFRHRI 58
||||||| |::| | |

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 17:22:14 ; Search time 61.48 Seconds
(without alignments)
643.173 Million cell updates/sec

Title: US-09-898-216-1

Perfect score: 1767
Sequence: 1 MLARAARGHMGFAEGSLSTG.....RDVQSTASXDELDLRVMS 356

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1763	99.8	356	19	AAW57232
2	1646	93.2	356	20	AAW78160
3	1646	93.2	356	22	AA892659
4	1646	93.2	378	22	ABBI1981
5	1646	93.2	378	22	AAW79550
6	1643	92.5	3007	22	AAU28053
7	1635	92.5	358	22	AAW78566
8	1493	84.5	306	20	AAW78284
9	1064.5	60.2	323	22	AB559541
10	1011	57.2	222	22	AAU33158
11	946	53.5	463	22	AAU28241

12	826	46.7	411	21	AA620661	Arabidopsis thalia
13	826	46.7	411	21	AA650184	Arabidopsis thalia
14	826	46.7	515	21	AA631395	Arabidopsis thalia
15	816.5	46.2	382	21	AA648817	Arabidopsis thalia
16	816.5	46.2	401	21	AA648816	Arabidopsis thalia
17	816.5	46.2	411	21	AA648815	Arabidopsis thalia
18	794.5	45.0	394	21	AA584663	Amino acid sequenc
19	716	40.5	315	21	AA585555	N. meningitidis am
20	716	40.5	315	21	AA585558	N. meningitidis am
21	716	40.5	315	21	AA525625	N. meningitidis am
22	716	40.5	315	21	AA525628	N. meningitidis am
23	716	40.5	315	21	AA574989	Neisseria meningit
24	716	40.5	315	21	AA574992	Neisseria meningit
25	716	40.5	315	21	AA575780	Neisseria meningit
26	716	40.5	315	21	AA575783	Neisseria meningit
27	716	40.5	315	21	AA575886	Neisseria strain z
28	716	40.5	315	21	AA575900	Neisseria strain z
29	715.5	40.5	315	21	AA575898	Neisseria strain z
30	714.5	40.4	315	21	AA575889	Neisseria strain z
31	714.5	40.4	315	21	AA575889	Neisseria strain z
32	714.5	40.4	315	21	AA575890	Neisseria strain z
33	714.5	40.4	315	21	AA575891	Neisseria strain z
34	714.5	40.4	315	21	AA575901	Neisseria strain z
35	712.5	40.3	315	21	AA585856	N. meningitidis am
36	712.5	40.3	315	21	AA585856	N. meningitidis am
37	712.5	40.3	315	21	AA574991	Neisseria meningit
38	712.5	40.3	315	21	AA575781	Neisseria meningit
39	712.5	40.3	315	21	AA575887	Neisseria strain z
40	712.5	40.3	315	21	AA575883	Neisseria strain z
41	712.5	40.3	315	21	AA575894	Neisseria strain z
42	712.5	40.3	315	21	AA575895	Neisseria strain z
43	712.5	40.3	315	21	AA575896	Neisseria strain z
44	712.5	40.3	315	21	AA575897	Neisseria strain z
45	712.5	40.3	315	21	AA575899	Neisseria strain z

ALIGNMENTS

RESULT 1	
ID	AAW57232 standard; Protein: 356 AA.
AC	AAW57232:
XX	XX
DT	03-AUG-1998 (first entry)
XX	XX
DE	Human integral membrane protein.
XX	XX
KW	Human; integral membrane protein; IMP; cancer; anaemia; prostate;
KW	breast; pancreatic; tumour; ion transport.
XX	XX
OS	Homo sapiens.
XX	XX
FH	Key
FT	Misc-difference 143
FT	/label="unknown
FT	/note="encoded by NTC"
FT	Misc-difference 173
FT	/note="encoded by CTN"
FT	Misc-difference 346
FT	/label="unknown
FT	/note="encoded by NTC"
XX	XX
PN	US5763589-A.
XX	XX
PD	09-JUN-1998.
XX	XX
PF	09-JAN-1997; 97US-0781562.
XX	XX
PR	09-JAN-1997; 97US-0781562.
XX	XX
PA	(INCY-) INCYTE PHARM INC.

XX 13-SEP-2001.
 PD 05-MAR-2001; 2001MO-US04942.
 PF 07-MAR-2000; 2000US-0519705.
 PR 19-MAY-2000; 2000US-0574454.
 PR 17-JUN-2000; 2000US-0596193.
 PR 14-JUL-2000; 2000US-0616847.
 PR 19-SEP-2000; 2000US-0665363.
 PR 20-OCT-2000; 2000US-0693267.
 XX
 PA (HYSE-) HYSEQ INC.
 XX Tang YN, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
 PI Zhao Qa, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AD, Wang J;
 DR WPI: 2001-589934/66.
 N-PSDB; AAS45141.
 XX
 PT Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis and treatment of
 PT cancer, neurological, inflammatory, and autoimmune disorders -
 XX
 PS Example 2; SEQ ID No 598; 107pp; English.
 XX
 CC The invention relates to novel isolated human secreted polypeptides (I)
 CC and polynucleotides (II). (I) and (II) are useful for treating
 CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
 CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
 CC involved in increasing haematopoiesis, stem cell survival, bone growth
 CC and remodeling. (I), (II) and modulators of (II) are useful for
 CC prophylaxis or treatment of one or more cancers. (II) is also useful for
 CC creating transgenic animals useful for studying the in vivo activities of
 CC the polypeptide as well as for studying modulators of the polypeptides.
 CC (I) induces the proliferation of neural cells and regeneration of nerve
 CC and brain tissue and is useful for the treatment of central and
 CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
 CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
 CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth, and in tissue repair, healing of burns, incisions,
 CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
 CC disorders, or periodontal disease. Furthermore, (I) is also useful for
 CC gut protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues, various immune deficiencies and
 CC disorders including severe combined immunodeficiency (SCID), bacterial or
 CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
 CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
 CC reactions and conditions, such as asthma or other respiratory problems.
 CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
 CC fertility, metabolism, catabolism, anabolism, storage or elimination of
 CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
 CC analgesic effects or other pain reducing effects, immunoglobulin like
 CC activity and can act as an antigen in a vaccine composition to raise an
 CC immune response. AAU28020-AAU28395 represent novel human secreted protein
 CC amino acid sequences of the invention.
 CC
 XX Sequence 463 AA:
 SQ

Query Match 14.0%; Score 50; DB 22; Length 463;
 Best Local Similarity 100.0%; Pred. No. 2.7e-40; Indels 0; Gaps 0;
 Matches 50; Conservative 0; Mismatches 0;

QY 27 RASSGLPRTNTVLEVPQGEAWVERMGREFRHLIEPGLNIIPLVDLRIRYV 76
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 108 reassgipntvtvlfpqgeawwvrcmrgfrhrliepglnllplvdrlriryv 157

RESULT 13

AAW78286
 ID AAW78286 standard: Protein; 27 AA.
 XX
 AC AAW78286;
 XX
 DT 13-APR-1999 (first entry)
 XX
 DE Fragment of human secreted protein encoded by gene 35.
 XX
 KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 XX
 OS Homo sapiens.
 XX
 PN W09856804-A1.
 XX
 PD 17-DEC-1998.
 XX
 PF 11-JUN-1998; 98MO-US12125.
 XX
 PR 02-OCT-1997; 97US-0061060.
 PR 13-JUN-1997; 97US-0049547.
 PR 13-JUN-1997; 97US-0049548.
 PR 13-JUN-1997; 97US-0049549.
 PR 13-JUN-1997; 97US-0049550.
 PR 13-JUN-1997; 97US-0049606.
 PR 13-JUN-1997; 97US-0049607.
 PR 13-JUN-1997; 97US-0049608.
 PR 13-JUN-1997; 97US-0049609.
 PR 13-JUN-1997; 97US-0049610.
 PR 13-JUN-1997; 97US-0049611.
 PR 13-JUN-1997; 97US-0050566.
 PR 13-JUN-1997; 97US-0050901.
 PR 13-JUN-1997; 97US-0052989.
 PR 08-JUL-1997; 97US-0051919.
 PR 18-AUG-1997; 97US-0055984.
 PR 12-SEP-1997; 97US-0058665.
 PR 12-SEP-1997; 97US-0058668.
 PR 12-SEP-1997; 97US-0058669.
 PR 12-SEP-1997; 97US-0058750.
 PR 12-SEP-1997; 97US-0058971.
 PR 12-SEP-1997; 97US-0058972.
 PR 12-SEP-1997; 97US-0058975.
 PR 02-OCT-1997; 97US-0060834.
 PR 02-OCT-1997; 97US-0060844.
 PR 02-OCT-1997; 97US-0060845.
 PR 02-OCT-1997; 97US-0060865.
 PR 02-OCT-1997; 97US-0061059.
 XX
 PA (HDMA-) HUMAN GENOME SCI INC.
 XX
 PI Brewer LA, Edner R, Ferris AM, Feng P, Greene JM, Lafleur DW;
 PI Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM, Shi Y, Young P;
 PI Yu GL;
 XX
 DR WPI: 1999-080881/07.
 DR N-PSDB; AAX04345.
 XX
 PT New isolated human genes and the secreted polypeptides they encode -
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders
 XX
 PS Disclosure: Page 35; 380pp; English.
 CC
 CC This sequence represents a fragment of a secreted human protein encoded
 CC by the nucleic acid molecule detailed in the descriptor line. The gene
 CC can be used to generate fusion proteins by linking to the gene to a

```

XX OS Homo sapiens.
XX OS WO200179449-A2.
XX PN 25-OCT-2001.
XX PD 16-APR-2001; 2001WO-US08656.
XX PF 18-APR-2000; 2000US-0552929.
XX PR 26-JAN-2001; 2001US-0770160.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR WPI: 2001-611725/70.
XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic
XX PT vaccination, testing and therapy -
XX PS Claim 20; Page 716; 765pp; English.
XX CC The invention relates to novel human secreted polypeptides. The
XX CC polypeptides and antibodies to the polypeptides are useful for
XX CC determining the presence of or predisposition to a disease associated
XX CC with altered levels of polypeptide. The polypeptides are also useful for
XX CC identifying agents (agonists and antagonists) that bind to them. Cells
XX CC expressing the proteins are useful for identifying a therapeutic agent
XX CC for use in treatment of a pathology related to aberrant expression or
XX CC physiological interactions of the polypeptide. Vectors comprising
XX CC the nucleic acids encoding the polypeptides and cells genetically
XX CC engineered to express them are also useful for producing the proteins.
XX CC The proteins are useful in genetic vaccination, testing and
XX CC therapy, and can be used as nutritional supplements. They may be used to
XX CC increase stem cell proliferation; to regulate haematopoiesis; and in
XX CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
XX CC immune suppression and/or stimulation; as anti-inflammatory agents; and
XX CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid
XX CC sequences of novel human secreted proteins of the invention.
XX SQ Sequence 222 AA;

Query Match 32.6%; Score 116; DB 22; Length 222;
Best Local Similarity 100.0%; Pred. No. 3.2e-105;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 27 RASSGLPRTVTVLEVPQDEAMVVERMRGRHLEPGNLILFVLDRIIRYVSLKEIVINV 86
DB 27 rassglprntvvlfpvqdeamvvermgrfhrllpgrlnlllpvldiriryvgsllkeivinv 86
OY 87 PEQSAVTDNNTLQIDGVLYLRIMDPKASGVDEDEYAVNQLAQTMRSELGKIS 142
DB 87 pegasavtdnntlvqlidgvlylrimdpykasygvedeyavnlqlatmrselgkis 142

RESULT 11
ID AA009538 standard; Protein: 92 AA.
XX AA009538;
XX AC AA009538;
XX DT 06-NOV-2001 (first entry)
XX DE Human polypeptide SEQ ID NO 23430.
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX KW nervous system disorders; arthritis; inflammation.
XX OS Homo sapiens.

```

```

XX OS WO200164835-A2.
XX PN 07-SEP-2001.
XX PD 26-FEB-2001; 2001WO-US04927.
XX PF 28-FEB-2000; 2000US-0515126.
XX PR 18-MAY-2000; 2000US-0577409.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR WPI: 2001-514838/56.
XX DR N-PSDB; AA189469.
XX PT Isolated nucleic acids and polypeptides, useful for preventing
XX PT diagnosing and treating e.g. leukaemia, inflammation and immune
XX PT disorders -
XX PS Claim 20; SEQ ID NO 23430; 1399pp + Sequence Listing; English.
XX CC The invention relates to human polynucleotides (AA179941-AA193641) and
XX CC the encoded proteins (AA000010-AA013910) that exhibit activity relating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 92 AA;

Query Match 14.6%; Score 52; DB 22; Length 92;
Best Local Similarity 100.0%; Pred. No. 6.9e-43;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 294 NTILPSPGQVTSWVAQAMGVYGLTKAPVPGTPDSISGSSRDVQGNDA 345
DB 30 ntillpspgdvtswvaqamgyygaltkapvpgrtpdsissgsrdvqgudas 81

RESULT 12
ID AAU28241 standard; Protein: 463 AA.
XX AAU28241;
XX AC AAU28241;
XX DT 18-DEC-2001 (first entry)
XX DE Novel human secretory protein, Seq ID NO 598.
XX KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
XX KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
XX KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
XX KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
XX KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
XX KW gut protection; lung; liver fibrosis; immune deficiency; infection;
XX KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
XX KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
XX KW fertility; analgesic; pain; antigen.
XX OS Homo sapiens.
XX PN WO200166689-A2.

```

PR 30-NOV-2000; 2000US-0728422.
XX (HYSE-) HYSEQ INC.
XX
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX WPI: 2001-476283/51.
DR N-PSDB: AAK51699.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
PS
PS Claim 20; Page 3486-3487; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAK78323-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation and which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAK80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
XX Sequence 358 AA:
SQ
Query Match 55.6%; Score 198; DB 22; Length 358;
Best Local Similarity 99.7%; Pred. No. 2.5e-185;
Matches 298; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 27 RASSGLPRTVTVLFVPOQAMVVERKGRFRHLEPGLNLLIFVLDRIRYVQSLKEIVNV 86
DB |||||||prvtvllfvpqgaavvermgtrfrlllepglllllpvldrlryvgalketvlnv 86
OY 87 PQGSNTLDNTLQIDGVLYLRIMDPYKASYGDEPETAVTOLAQTTMSELGKLSXDKV 146
DB |||||||ntldntlqidgvlylrimdpkasygvedpetavtolaqttmselgklsidkv 146
OY 87 pegasvclndnvlqldgvylyrlmipykasgyvedpeyavtlqqltmselgklsidkv 146
DB |||||||clndnvlqldgvylyrlmipykasgyvedpeyavtlqqltmselgklsidkv 146
OY 147 FFRERSLNASTYDAINQADCGICRLREIKDIHVPVFKSMQVFAERRKRTATVE 206
DB |||||||freslnastydainqadcgicrlreikdihvpvfksmqvfaerrkrattve 206
OY 147 ftereslnastydainqadcgicrlreikdihvpvfksmqvfaerrkrattve 206
DB |||||||tereslnastydainqadcgicrlreikdihvpvfksmqvfaerrkrattve 206
OY 207 SEGTRESAIVNAEGKKQAQIILASEAEKAEQINQAGASAVLAKAKAKAEAIRIILAAALT 266
DB |||||||segtresainvaeqkkaqiilaseaekeqinqagasaavlakakakaeariilaalt 266
OY 207 segtresainvaeqkkaqiilaseaekeqinqagasaavlakakakaeariilaalt 266
DB |||||||segtresainvaeqkkaqiilaseaekeqinqagasaavlakakakaeariilaalt 266
OY 267 OHNGGAAASLTVAEOYVSFSLAKDSNTILPSNPGDVTSMVAQAMGYGALTAPVP 325
DB |||||||ohnggaaasltvaeoyvsfslakdsntilpsnpgdvtsmvaqamgygaltapvp 325
OY 267 qingdaaasltvaeogyvsfslakdsntilpsnpgdvtsmvaqamgygaltapvp 325
DB |||||||qingdaaasltvaeogyvsfslakdsntilpsnpgdvtsmvaqamgygaltapvp 325

RESULT 9
AAU33159
ID AAU33159 standard; Protein: 2797 AA.
XX
XX AAU33159;
AC
AC 18-DEC-2001 (first entry)
DT
DT 18-DEC-2001 (first entry)
DE
DE Novel human secreted protein #3650.
XX
XX Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW Immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX

OS Homo sapiens.
XX
XX PN WO200179449-A2.
XX
XX PD 25-OCT-2001.
XX
XX PF 16-APR-2001; 2001WO-US08656.
XX
XX PR 18-APR-2000; 2000US-0552929.
XX PR 26-JAN-2001; 2001US-0770160.
XX
XX PA (HYSE-) HYSEQ INC.
XX
XX PI Tang YT, Liu C, Drmanac RT;
XX
XX WPI: 2001-611725/70.
DR
DR Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -
PT
PT Claim 20; Page 716; 765pp; English.
XX
XX The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC Immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.
XX
XX Sequence 2797 AA:
SQ
Query Match 34.6%; Score 123; DB 22; Length 2797;
Best Local Similarity 100.0%; Pred. No. 4.5e-111;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 223 QAOIILASEAEKAEQINQAGASAVLAKAKAKAEAIRIILAAALTQHNGDAASLTVAEY 282
DB |||||||aoiilaseaekeqinqagasaavlakakakaeariilaaltqhngdaasltvaeqy 282
OY 283 VSAFSLAKDSNTILPSNPGDVTSMVAQAMGYGALTAPVPGTPDSLSSGSSRPDOCT 342
DB |||||||safslakdsntilpsnpgdvtsmvaqamgygaltapvpdpdsssgssrdovgt 342
OY 61 vsafslakdsntilpsnpgdvtsmvaqamgygaltapvpdpdsssgssrdovgt 342
DB |||||||vsafslakdsntilpsnpgdvtsmvaqamgygaltapvpdpdsssgssrdovgt 342
OY 343 DAS 345
DB |||
DB 121 das 123

RESULT 10
AAU33158
ID AAU33158 standard; Protein: 222 AA.
XX
XX AAU33158;
AC
AC 18-DEC-2001 (first entry)
DT
DT 18-DEC-2001 (first entry)
DE
DE Novel human secreted protein #3649.
XX
XX Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW Immune suppression; immune stimulation; anti-inflammatory; leukaemia.
KW

OY 327 TPDSLSGSSRDVQGTDA 345
 ||||||||||||||||
 Db 327 tpdslssgssrdvqgtas 345

RESULT 7
 AAM78284
 AAM78284 standard; Protein: 306 AA.
 XX
 AC AAM78284;
 XX
 DT 13-APR-1999 (first entry)
 XX
 DE Fragment of human secreted protein encoded by gene 35.
 XX
 KW Human: secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
 XX
 OS Homo sapiens.
 XX
 PN MO9856804-A1.
 XX
 PD 17-DEC-1998.
 XX
 PF 11-JUN-1998; 98WO-US12125.
 XX
 PR 02-OCT-1997; 97US-0061060.
 PR 13-JUN-1997; 97US-0049547.
 PR 13-JUN-1997; 97US-0049548.
 PR 13-JUN-1997; 97US-0049549.
 PR 13-JUN-1997; 97US-0049550.
 PR 13-JUN-1997; 97US-0049606.
 PR 13-JUN-1997; 97US-0049607.
 PR 13-JUN-1997; 97US-0049608.
 PR 13-JUN-1997; 97US-0049609.
 PR 13-JUN-1997; 97US-0049610.
 PR 13-JUN-1997; 97US-0049611.
 PR 13-JUN-1997; 97US-0050566.
 PR 13-JUN-1997; 97US-0050901.
 PR 13-JUN-1997; 97US-0052989.
 PR 08-JUL-1997; 97US-0051919.
 PR 18-AUG-1997; 97US-0055984.
 PR 12-SEP-1997; 97US-0058665.
 PR 12-SEP-1997; 97US-0058666.
 PR 12-SEP-1997; 97US-0058669.
 PR 12-SEP-1997; 97US-0058750.
 PR 12-SEP-1997; 97US-0058971.
 PR 12-SEP-1997; 97US-0058972.
 PR 12-SEP-1997; 97US-0058975.
 PR 02-OCT-1997; 97US-0060834.
 PR 02-OCT-1997; 97US-0060841.
 PR 02-OCT-1997; 97US-0060844.
 PR 02-OCT-1997; 97US-0060865.
 PR 02-OCT-1997; 97US-0061059.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Brewer LA, Ebner R, Ferlie AM, Feng P, Greene JM, Lafleur DM;
 PI Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM, Shi Y, Young P;
 PI Yu GL;
 XX
 DR WPI: 1999-080881/07.
 DR N-PSDB: AAX04345.
 XX
 PT New isolated human genes and the secreted polypeptides they encode -
 useful for diagnosis and treatment of e.g. cancers, neurological

PT disorders, immune diseases, inflammation or blood disorders
 XX
 PS Disclosure: Page 35; 380pp; English.
 XX
 CC This sequence represents a fragment of a secreted human protein encoded
 CC by the nucleic acid molecule detailed in the descriptor line. The gene
 CC can be used to generate fusion proteins by linking to the gene to a
 CC human immunoglobulin Fc portion (e.g. AAX04302) for increasing the
 CC stability of the fused protein as compared to the human protein only.
 CC The invention relates to 86 novel genes and their fragments (nucleic
 CC acid sequences: AAX04311-X04410; amino acid sequences AAM78126-W78225)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 86
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AAX04311 for described uses).
 CC
 CC Sequence 306 AA:
 XX
 SQ

Query Match 56.7%; Score 202; DB 20; Length 306;
 Best Local Similarity 100.0%; Pred. No. 2; 6e-189;
 Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 144 DKVFERESLNASIVDAINQADCGICRLREIKDIHVPPVKESMOMOVEAERKRAT 203
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 94 dkvtrereslnasivdaingdaedcwqjrlrlyedlnppvksmqmgyaeetrkat 153

OY 204 VLESEGTRESAINVAEGKKQAOIILASEAKEOINQAGASAVLAKAKAKAIAIRILAA 263
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 154 vlesegtresainvnaegkkgqqlaseaekeqinqaageasaavilakakakeairilaa 213

OY 264 ALTOHNGDAASLTVAEOYVSFSLKADSNITILPSPNGVTSNVAQAMGYGALTAP 323
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 214 altqhngdaaasltvaeqyvsafsklaxdsnllilpsnpgdvtsmvdaqmgygaltkap 273

OY 324 VPGTPDLSGSSRDVQGTDA 345
 ||||||||||||||||
 Db 274 vpgtpdlsissgssrdvqgtas 295

RESULT 8
 AAM78566
 AAM78566 standard; Protein: 358 AA.
 XX
 AC AAM78566;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human protein SEQ ID NO 1228.
 XX
 KW Human: cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; hematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US04098.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.

CC Inflammation.

CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 (AAH80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication.

XX Sequence 378 AA:

Query Match 61.2%; Score 218; DB 22; Length 378;
Best Local Similarity 99.7%; Pred. No. 6.9e-205;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 27 RASSGLPRNTVVLFPVQOEAWVERMGRRHRLLEPGLNLIIPVLDRIYVOSLKEIVINW 86
DB 49 rassg1prntvvlfpvgqawvermgrrhrllepglnlipvlidrlirygslkeivinv 108
OY 87 PEQSAVTLDNWTLQIDGVLYLRIMDPYKASGYVEDEPEYAVNOLAQTMRSEIGKLSXKV 146
DB 109 pegsavtldnwtlqidgvylylrimdpykasyvedpeyavncqlaqtcmseigklsldkv 168
OY 147 FRERESLNASIVDAINQADACGIRCLRYEIKDIHVPRVKSQMOMQVEARRKRATYLE 206
DB 169 frereslnasivdaingadacgirclyelkdihvprvksmqmvyaeerrkratyle 228
OY 207 SEGTRESAINVAEGKKQAOIILASEAKAEOINQAGASAVILAKAKAKAEAIRIIAAALT 266
DB 229 segtresainvaeqgkkqgqilaseeakeqinqaagesavilakakakaearriiiaaalt 288
OY 267 OHNGDAASLTVAEOYVAFSKLAKDSNTIILPSNPGDVTSMVAAMGVGALTAPVPG 326
DB 289 ohngdaasltvaeqyvaafsklakdsntllipnpgdvtsmvagamvygaltkavpg 348
OY 327 TPDSLSSGSSRDVQCTDAS 345
DB 349 tpdslssgsrdvgqtdas 367

RESULT 6
AAU28053
ID AAU28053 standard; Protein: 3007 AA.
XX
AC AAU28053;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secretory protein, Seq ID No 222.
XX
KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock; ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy; transgenic animal; Alzheimer's disease; Parkinson's disease; burn; amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia; ulcer; osteoporosis; bone degenerative disorder; periodontal disease; gut protection; lung; liver fibrosis; immune deficiency; infection; severe combined immunodeficiency; SCID; autoimmune disorder; allergy; multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma; fertility; analgesic; pain; antigen.
KW
KW Homo sapiens.
OS
XX
PN MO200166689-A2.
XX
PD 13-SEP-2001.
XX
XX 05-MAR-2001; 2001WO-US04942.
XX
XX 07-MAR-2000; 2000US-0519705.
PR 19-MAY-2000; 2000US-0574454.
PR 17-JUN-2000; 2000US-0596193.
PR 14-JUL-2000; 2000US-0616847.
PR 19-SEP-2000; 2000US-0665363.
PR 20-OCT-2000; 2000US-0693267.
XX
XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
PI Zhao Qa, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
XX
XX WPI: 2001-589934/66.
DR N-PSDB: AAS44953.
XX

PT Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis and treatment of
PT cancer, neurological, inflammatory, and autoimmune disorders -
XX
PS Example 3: SEQ ID No 222: 107pp; English.

XX The invention relates to novel isolated human secreted polypeptides (I)
CC and polynucleotides (II). (I) and (II) are useful for treating
CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
CC involved in increasing haematopoiesis, stem cell survival, bone growth
CC and remodeling. (I), (II) and modulators of (II) are useful for
CC prophylaxis or treatment of one or more cancers. (II) is also useful for
CC creating transgenic animals useful for studying the in vivo activities of
CC the polypeptide as well as for studying modulators of the polypeptides.
CC (I) induces the proliferation of neural cells and regeneration of nerve
CC and brain tissue and is useful for the treatment of central and
CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
CC activity, regulation of haematopoiesis and is useful for treating myeloid
CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
CC tissue growth, and in tissue repair, healing of burns, incisions,
CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
CC disorders, or periodontal disease. Furthermore, (I) is also useful for
CC gut protection or regeneration and treatment of lung or liver fibrosis.
CC Reperfusion injury in various tissues, various immune deficiencies and
CC disorders including severe combined immunodeficiency (SCID), bacterial or
CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
CC reactions and conditions, such as asthma or other respiratory problems.
CC In addition, (I) affects biorhythms or circadian cycles of rhythms.
CC fertility, metabolism, catabolism, anabolism, storage or elimination of
CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
CC analgesic effects or other pain reducing effects, immunoglobulin like
CC activity and can act as an antigen in a vaccine composition to raise an
CC immune response. AAU28020-AAU28395 represent novel human secreted protein
CC amino acid sequences of the invention.
XX

Seq Sequence 3007 AA:

Query Match 61.2%; Score 218; DB 22; Length 3007;
Best Local Similarity 99.7%; Pred. No. 4.5e-204;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 27 RASSGLPRNTVVLFPVQOEAWVERMGRRHRLLEPGLNLIIPVLDRIYVOSLKEIVINW 86
DB 27 rassg1prntvvlfpvgqawvermgrrhrllepglnlipvlidrlirygslkeivinv 86
OY 87 PEQSAVTLDNWTLQIDGVLYLRIMDPYKASGYVEDEPEYAVNOLAQTMRSEIGKLSXKV 146
DB 87 pegsavtldnwtlqidgvylylrimdpykasyvedpeyavncqlaqtcmseigklsldkv 146
OY 147 FRERESLNASIVDAINQADACGIRCLRYEIKDIHVPRVKSQMOMQVEARRKRATYLE 206
DB 147 frereslnasivdaingadacgirclyelkdihvprvksmqmvyaeerrkratyle 206
OY 207 SEGTRESAINVAEGKKQAOIILASEAKAEOINQAGASAVILAKAKAKAEAIRIIAAALT 266
DB 207 segtresainvaeqgkkqgqilaseeakeqinqaagesavilakakakaearriiiaaalt 266
OY 267 OHNGDAASLTVAEOYVAFSKLAKDSNTIILPSNPGDVTSMVAAMGVGALTAPVPG 326
DB 267 ohngdaasltvaeqyvaafsklakdsntllipnpgdvtsmvagamvygaltkavpg 326

PD 09-AUG-2001.
 XX
 PF 05-FEB-2001: 2001MO-US03800.
 XX
 PR 03-FEB-2000: 2000US-0496914.
 PR 27-APR-2000: 2000US-0560875.
 XX
 PA (HYSE-) HYSEQ INC.
 PI Tang YT, Liu C, Drmanac RT.
 XX
 DR MPI: 2001-457740/49.
 DR N-PSDB: ABA09225.
 XX
 PT Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer -
 XX
 XX Claim 20: Page 291: 1963pp: English.
 XX
 PS Sequences ABA10981-AB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC hematopoiesis regulatory activity; tissue growth factor activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention.
 XX
 XX Sequence 378 AA:
 S0

Query Match 61.2%: Score 218: DB 22: Length 378:
 Best Local Similarity 99.7%: Pred. No. 6.9e-205:
 Matches 318: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

0y 27 RASGCLPNTVAVLPYQOEAHVVERMGRFRLLEGLNLLPVDLRIRYVSLKELIVNV 86
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 49 rassgjlptntvvlfpvgqeaavvermgfrfhrllpglllllpvldrlryvgsllkeltivn 108
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 0y 87 PEQSAVLTNDNTLTQIDGLVLYRIMDPKASGVGVEDPEYAVQLOATWRSELGKISXKV 146
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 109 pegsavltndnvtlqldgvllylrilmopykasygvedpeyavqlatqumtse.gklsldkv 168

0y 147 FRERESLNASTVDAINQADQWGIKRLRYEIKDIIHVPKPKESNMQOVKERRRRATVLE 206
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 169 frereslnastvdaingadqwgirclryelkldihvprvkemqmqveeerikratvle 228
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 0y 207 SEGTRESAINVAEKKKQAOILLASFAEKAKEQINQAGFSAVLLAKAKKAEIRLLAALT 266
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 229 segtresainvaeqkkqagillaseaeakeqingaaesavllakakaeairllaaalt 288
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 0y 267 OHNGDAASLTVAEQYVSAFSLKAKDSNTILLPSNPGDVTSMVAQAMGVGALTAKAPVPG 326
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 289 qhngdaaasltvaeqyvsafslkaksdntllpsnpgdvtsmvgaqangvygalckapvpg 348
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 0y 327 TPDSLSSGSSRDVGTDPAS 345
 ||||||||||||||||||||
 Db 349 tpdslssgsrdvgtgdas 367
 ||||||||||||||||||||

RESULT 5
 AAM79550 standard; Protein: 378 AA.
 ID AAM79550:
 XX
 AC AAM79550:
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human protein SEQ ID NO 3196.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; hematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukemia;
 KW nervous system disorder; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN MO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001: 2001MO-US04098.
 XX
 PR 03-FEB-2000: 2000US-0496914.
 PR 27-APR-2000: 2000US-0560875.
 PR 20-JUN-2000: 2000US-0598075.
 PR 19-JUL-2000: 2000US-0620325.
 PR 15-SEP-2000: 2000US-0654936.
 PR 01-SEP-2000: 2000US-0663561.
 PR 20-OCT-2000: 2000US-0693325.
 PR 30-NOV-2000: 2000US-0728422.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y,
 PI Zhao QH, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 DR MPI: 2001-476283/51.
 DR N-PSDB: AAK52683.
 XX
 PT Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 XX
 PS Claim 20: Page 281-282: 6221pp: English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAW80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, hematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukemia, nervous system disorders, arthritis and

Db 27 rassg|prntv|lfvpggawvvermg|fhrll|epg|nll|pvidr|lyvgs|lkelv|nv 86
 Oy 87 PEQSAVTLDNVTLQIDGVLYLRIMDPYKASYGVEDPEYAVTQLAQTTMRSELGKLSXDKV 146
 Db 87 pegasv|tldn|v|l|q|d|g|v|l|y|l|r|m|d|p|k|a|s|y|v|e|d|p|e|y|a|v|t|q|l|a|g|t|m|s|e|l|g|k|l|s|d|k|v 146
 Oy 147 FRERESLNASTVDAINQADCGTICLREYETKDIHVPRVRESKMOVQEAERKRAVLE 206
 Db 147 f|r|e|r|e|s|l|n|a|s|t|v|d|a|i|n|q|a|d|c|g|t|i|c|l|r|e|y|e|t|k|d|i|h|v|p|r|v|r|e|s|m|q|v|e|a|e|r|k|r|a|v|l|e 206
 Oy 207 SEGTRESAINVAEGKKQAOILASEAEKAEQINQAGEASAVLAKAKAEAIRIRIAALT 266
 Db 207 s|e|g|t|r|e|s|a|i|n|v|a|e|g|k|k|q|a|l|a|s|e|e|k|e|q|i|n|q|a|g|e|a|s|a|v|l|a|k|a|k|a|e|a|i|r|i|a|a|l|t 266
 Oy 267 QHNGDAASLTVAEQYVSAFSLAKDSNTILLPSNPGDVTSMVQAMGVGALTAKAPVP 326
 Db 267 q|h|n|g|d|a|a|s|l|t|v|a|e|q|y|v|s|a|f|s|l|a|k|d|s|n|t|i|l|l|p|s|n|p|g|d|v|t|s|m|v|q|a|m|g|v|g|a|l|t|a|k|a|p|p|g 326
 Oy 327 TPDSLSGSSRDVQCTDAS 345
 Db 327 t|p|d|s|l|s|g|s|s|r|d|v|q|c|t|d|a|s 345
 RESULT 3
 AAB92659
 ID AAB92659 standard; Protein: 356 AA.
 AC AAB92659;
 XX 26-JUN-2001 (first entry)
 DT 26-JUN-2001 (first entry)
 XX Human protein sequence SEQ ID NO:11010.
 DE Human protein sequence SEQ ID NO:11010.
 KW Human; primer: detection; diagnosis; antisense therapy; gene therapy.
 OS Homo sapiens.
 XX Homo sapiens.
 XX Ep1074617-A2.
 PD 07-FEB-2001.
 PF 28-JUL-2000: 2000EP-0116126.
 XX 28-JUL-1999: 99JP-0248026.
 PR 27-AUG-1999: 99JP-0300253.
 PR 11-JAN-2000: 2000JP-0118776.
 PR 02-MAY-2000: 2000JP-0183767.
 PR 09-JUN-2000: 2000JP-0241899.
 XX (HELI-) HELIX RES INST.
 PA Ota T, Sugiyama T, Nishikawa T, Hayashi K, Salto K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI: 2001-318749/34.
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX Claim 8: SEQ ID 11010; 2537pp + CD-ROM; English.
 PS The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX Sequence 356 AA:
 SQ
 Query Match 61.2%; Score 218; DB 22; Length 356;
 Best Local Similarity 99.7%; Pred. No. 6.5e-205;
 Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 27 RASSGLPRNTVLFVPGGEAWVVERMGFRHRLLEPGLNLLPVDLRIRYQSLKEIVNV 86
 Db 27 rassg|prntv|lfvpggawvvermg|fhrll|epg|nll|pvidr|lyvgs|lkelv|nv 86
 Oy 87 PEQSAVTLDNVTLQIDGVLYLRIMDPYKASYGVEDPEYAVTQLAQTTMRSELGKLSXDKV 146
 Db 87 pegasv|tldn|v|l|q|d|g|v|l|y|l|r|m|d|p|k|a|s|y|v|e|d|p|e|y|a|v|t|q|l|a|g|t|m|s|e|l|g|k|l|s|d|k|v 146
 Oy 147 FRERESLNASTVDAINQADCGTICLREYETKDIHVPRVRESKMOVQEAERKRAVLE 206
 Db 147 f|r|e|r|e|s|l|n|a|s|t|v|d|a|i|n|q|a|d|c|g|t|i|c|l|r|e|y|e|t|k|d|i|h|v|p|r|v|r|e|s|m|q|v|e|a|e|r|k|r|a|v|l|e 206
 Oy 207 SEGTRESAINVAEGKKQAOILASEAEKAEQINQAGEASAVLAKAKAEAIRIRIAALT 266
 Db 207 s|e|g|t|r|e|s|a|i|n|v|a|e|g|k|k|q|a|l|a|s|e|e|k|e|q|i|n|q|a|g|e|a|s|a|v|l|a|k|a|k|a|e|a|i|r|i|a|a|l|t 266
 Oy 267 QHNGDAASLTVAEQYVSAFSLAKDSNTILLPSNPGDVTSMVQAMGVGALTAKAPVP 326
 Db 267 q|h|n|g|d|a|a|s|l|t|v|a|e|q|y|v|s|a|f|s|l|a|k|d|s|n|t|i|l|l|p|s|n|p|g|d|v|t|s|m|v|q|a|m|g|v|g|a|l|t|a|k|a|p|p|g 326
 Oy 327 TPDSLSGSSRDVQCTDAS 345
 Db 327 t|p|d|s|l|s|g|s|s|r|d|v|q|c|t|d|a|s 345
 RESULT 4
 ABB11981
 ID ABB11981 standard; peptide: 378 AA.
 AC ABB11981;
 XX 11-JAN-2002 (first entry)
 DT 11-JAN-2002 (first entry)
 XX Human SLP-2 homologue, SEQ ID NO:2351.
 DE Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytosolic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnary; antitumor.
 XX Homo sapiens.
 OS Homo sapiens.
 XX WO200157188-A2.
 XX

XX Goli SK, Hillman JL;
 PI
 XX
 DR WPI: 1998-347418/30.
 N-PSDB: AAV28867.
 XX
 PT DNA encoding human integral membrane protein - useful for producing
 recombiant protein, for treatment of anaemia and cancer
 XX
 PS Claim 1; Column 35-38; 33pp; English.
 XX
 CC The present sequence represents human integral membrane protein (IMP).
 CC IMP may be administered to a subject to treat disorders associated
 CC with abnormal ion transport or membrane conductance as well as a
 CC variety of tumours, e.g. haemolytic anaemias and prostate, breast and
 CC pancreatic tumours. A vector capable of expressing IMP, or a fragment
 CC or a derivative thereof, may also be administered to a subject to treat
 CC the haemolytic anaemias and prostate, breast and pancreatic tumours.
 CC
 XX Sequence 356 AA:

Query Match 99.4%; Score 354; DB 19; Length 356;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLARAGHMGPEAGLSTGFWPSCGRASGLPNTVLEVPQOEAWVERMGRFRHILE 60
 |
 Db 1 mlaraaghhmgpfeaglstgfwpsgrassglpntvvlftpgqeaavvermgfrhille 60
 QY 61 PGLNLIPLVDLRIRYVSLKEIVINPEQSAVTLDNVTLQIDGVLYLRIMDPKASYGVE 120
 |
 Db 61 pglnlilpvlldriryvsllkeivinvpeqsaavlndnvtlqldgyllylrindpkyasygve 120
 QY 121 DPEAVVQLOAQOTMRSELGKSKXKVRERESLNASTYDAINQADCGITCLAREIKDI 180
 |
 Db 121 dpeavvqlaqotmrselegksskxkvrrereslnastvdaingadcgilrclyelkdi 180
 QY 181 HVPFRVESMQQVEAREKRRATVLESEGTRESAINVAEKKQKQILASEAKAQINO 240
 |
 Db 181 hvprvvesmqqqvearekrratvlesegtrresainvaeqgkqgillaseekqinqa 240
 QY 241 AGEASAVLAKKAKAEAIRILAAALTOHNGDAASLVYAEQVYSAFESKLADSNITLIPS 300
 |
 Db 241 ageasavlakakaeairilaaaltqhngdaaaslvyaegyafeskladsnitllips 300
 QY 301 NPGDVTGVVAQAMGVGALTAKAPVGPDPDSLSGSSRPVQGTDA SXDEELRVKMS 356
 |
 Db 301 npgdvtsmvgaqamvgvgtalkapvpdpdsissgssrdvgtddasxdeelrvkms 356

RESULT 2
 AAW78160
 ID AAW78160 standard; Protein: 356 AA.
 XX
 AC AAW78160;
 XX
 DT 13-APR-1999 (first entry)
 XX
 DE Human secreted protein encoded by gene 35 clone HTXCS21.
 XX
 XX Human: secreted protein: fusion protein: gene therapy: protein therapy;
 KW diagnosis: tissue: cancer: tumour: neurodegenerative disorder: leukaemia;
 KW developmental abnormality: foetal deficiency: blood: allergy: renal;
 KW immune system: asthma; lymphocytic disease: brain: hepatic; lymphoma;
 KW inflammation: ischaemic shock: Alzheimer's disease: restenosis: AIDS;
 KW cognitive disorder: schizophrenia; prostate: obesity: osteoclast; thymus;
 KW osteoporosis: arthritis; testis; lung; thyroiditis; thyroid: digestion;
 KW endocrine: metabolism; regulation; malabsorption; gastritis; neoplasm.
 XX
 XX Homo sapiens.
 OS
 XX
 PN M09856804-A1.

XX
 PD 17-DEC-1998.
 XX
 XX 11-JUN-1998; 98MO-US12125.
 XX
 PR 02-OCT-1997; 97US-0061060.
 PR 13-JUN-1997; 97US-0049547.
 PR 13-JUN-1997; 97US-0049548.
 PR 13-JUN-1997; 97US-0049549.
 PR 13-JUN-1997; 97US-0049550.
 PR 13-JUN-1997; 97US-0049606.
 PR 13-JUN-1997; 97US-0049607.
 PR 13-JUN-1997; 97US-0049608.
 PR 13-JUN-1997; 97US-0049609.
 PR 13-JUN-1997; 97US-0049610.
 PR 13-JUN-1997; 97US-0049611.
 PR 13-JUN-1997; 97US-0050566.
 PR 13-JUN-1997; 97US-0050901.
 PR 13-JUN-1997; 97US-0052989.
 PR 08-JUL-1997; 97US-0051919.
 PR 18-AUG-1997; 97US-0055984.
 PR 12-SEP-1997; 97US-0058665.
 PR 12-SEP-1997; 97US-0058668.
 PR 12-SEP-1997; 97US-0058669.
 PR 12-SEP-1997; 97US-0058750.
 PR 12-SEP-1997; 97US-0058971.
 PR 12-SEP-1997; 97US-0058972.
 PR 12-SEP-1997; 97US-0058975.
 PR 02-OCT-1997; 97US-0060834.
 PR 02-OCT-1997; 97US-0060841.
 PR 02-OCT-1997; 97US-0060844.
 PR 02-OCT-1997; 97US-0060865.
 PR 02-OCT-1997; 97US-0061059.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Brewer LA, Ebner R, Ferrie AM, Feng P, Greene JM, Lafleur DW;
 PI Moore PA, NI J, Olsen HS, Rosen CA, Ruben SM, Shi Y, Young P;
 PI Yu GL;
 DR WPI: 1999-080881/07.
 N-PSDB: AAX04345.
 XX
 PT New isolated human genes and the secreted polypeptides they encode -
 useful for diagnosis and treatment of e.g. cancers, neurological
 disorders, immune diseases, inflammation or blood disorders
 XX
 PS Claim 11; Page 282-283; 380pp; English.
 XX
 CC This sequence represents a secreted human protein encoded by the gene
 CC clone detailed in the descriptor line. The gene can be used to generate
 CC fusion proteins by linking to the gene to a human immunoglobulin Fc
 CC portion (e.g. AAX04302) for increasing the stability of the fused
 CC protein as compared to the human protein only.
 CC The invention relates to 86 novel genes and their fragments (nucleic
 CC acid sequences: AAX04311-04410; amino acid sequences AAW78126-W8225)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 86
 CC (see AAX04311 for described uses).
 CC
 XX Sequence 356 AA:

Query Match 61.2%; Score 218; DB 20; Length 356;
 Best Local Similarity 99.7%; Pred. No. 6; 56-205;
 Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 27 RASSGLPNTVVLFPQOEAWVERMGRFRHILEGLNLIPLVDLRIRYVSLKEIVNV 86
 |
 |

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

```
Run on:      September 22, 2002, 18:22:47 ; Search time 31.2 Seconds
              (without alignments)
              1267.380 Million cell updates/sec
```

Title: US-09-898-216-1

Sequence: 1 MLARAARGHWGPF~~AEGL~~STG.....RDVQGT~~DASXDE~~ELDRVKMS 356

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size

Total number of hits satisfying chosen parameters: 747574

```
Minimum DB seq length: 0
```

Maximum DB seq length: 20000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_032802:*

1:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1980.DAT*
2:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1981.DAT*
3:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1982.DAT*
4:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1983.DAT*
5:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1984.DAT*
6:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1985.DAT*
7:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1986.DAT*
8:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1987.DAT*
9:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1988.DAT*
10:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1989.DAT*
11:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1990.DAT*
12:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1991.DAT*
13:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1992.DAT*
14:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1993.DAT*
15:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1994.DAT*
16:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1995.DAT*
17:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1996.DAT*
18:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1997.DAT*
19:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1998.DAT*
20:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA1999.DAT*
21:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA2000.DAT*
22:	/SIDS1/gcgdata/genseq/genseqp-emb1/AA2001.DAT*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	354	99.4	356	19	AAW57232	Human integral mem
2	218	61.2	356	20	AAW76160	Human secreted prot
3	218	61.2	356	22	AAB92659	Human protein sequ
4	218	61.2	378	22	ABB11981	Human SLP-2 homold
5	218	61.2	378	22	AAW79550	Human protein SEQ
6	218	61.2	3007	22	AAU28053	Novel human secre
7	202	56.7	306	20	AAW78284	Fragment of human
8	198	55.6	358	22	AAW78566	Human protein SEQ
9	123	34.6	2797	22	AAU31159	Novel human secre
10	116	32.6	222	22	AAU31158	Novel human secre
11	52	14.6	92	22	AAO09538	Human polypeptide

12	50	4.0	463	22	AAU78241	Novel human secret
13	27	7.6	27	20	AAW78286	Fragment of human
14	26	7.3	26	20	AAW78285	Fragment of human
15	26	7.3	26	20	AAW78287	Fragment of human
16	24	6.7	70	20	AAW78288	Fragment of human
17	20	5.6	323	22	ABBS59541	Drosophila melanog
18	13	3.7	100	21	AAAG50216	Arabidopsis thallia
19	13	3.7	339	21	AAAG34300	Arabidopsis thallia
20	13	3.7	358	21	AAAG42899	Arabidopsis thallia
21	13	3.7	369	21	AAAG42988	Arabidopsis thallia
22	13	3.7	382	21	AAAG48817	Arabidopsis thallia
23	13	3.7	394	21	AAAG4663	Amino acid sequenc
24	13	3.7	401	21	AAAG48816	Arabidopsis thallia
25	13	3.7	411	21	AAAG30661	Arabidopsis thallia
26	13	3.7	411	21	AAAG48815	Arabidopsis thallia
27	13	3.7	411	21	AAAG50184	Arabidopsis thallia
28	13	3.7	515	21	AAAG31395	Arabidopsis thallia
29	10	2.8	10	22	AAU257867	Breast cancer-asso
30	8	2.2	15	22	AAU25787	Breast cancer-asso
31	8	2.2	43	22	ABB27904	Human peptide #555
32	8	2.2	43	22	ABB33076	Peptide #562 encod
33	8	2.2	43	22	ABBI8544	Protein #543 encod
34	8	2.2	43	22	AAAG53873	Human brain expres
35	8	2.2	43	22	AAAG66261	Human bone marrow
36	8	2.2	43	22	AAAI4131	Peptide #565 encod
37	8	2.2	43	22	AAAG26537	Peptide #574 encod
38	8	2.2	43	22	AAAO1869	Peptide #551 encod
39	8	2.2	304	22	AAAG91264	C glutimicm cul prote
40	8	2.2	554	22	ABAB71946	Drosophila melanog
41	8	2.2	1015	22	ABAG04137	Novel human diagno
42	8	2.2	1443	22	ABBB68472	Drosophila melanog
43	7	2.0	14	22	AAAB8061	Human peptide #133
44	7	2.0	14	22	AAAB8061	Human peptide #133
45	7	2.0	16	18	AAAB50219	KAE16, amphiphilic

ALIGNMENTS

RESULT	1
AAW57232	ID
AAW57232	standard; Protein; 356 AA.
XX	XX
AC	AAW57232;
XX	03-AUG-1998 (first entry)
DT	XX
XX	Human integral membrane protein.
DE	XX
Human; integral membrane protein; IMP; cancer: anaemia; prostate;	XX
breast; pancreatic; tumour; ion transport.	XX
XX	OS
Homo sapiens.	XX
Key	Location/Qualifiers
FH	Key
FT	Misc-difference 143
FT	/label= unknown
FT	/note= "encoded by NTG"
FT	Misc-difference 173
FT	/note= "encoded by CTN"
FT	Misc-difference 346
FT	/label= unknown
FT	/note= "encoded by NTG"
XX	XX
US5763589-A.	XX
XX	09-JUN-1998.
PD	XX
09-JAN-1997;	97US-0781562.
PE	XX
09-JAN-1997;	97US-0781562.
PR	XX
XX	XX
PA	(INCY-) INCYTE PHARM INC.

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL035602; CAB8270.1; -
 DR EMBL: AL161571; CAB81408.1; -
 DR InterPro: IPR001107; Band_7.
 DR InterPro: IPR001972; Stomatin.
 DR Pfam: PF01145; Band_7; 1.
 DR PRINTS: PR00721; STOMATIN.
 DR SMART: SM00244; PHB; 1.
 DR Hypothetical protein.
 KW Hypothetical protein.
 SQ SEQUENCE 515 AA; 55923 MW; 420FE3DB5B7AFDBA CRC64;

Query Match 3.1%; Score 13; DB 10; Length 515;
 Best Local Similarity 100.0%; Pred. No. 0.00047;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 128 TOLAOTMRSELC 140
 DB 153 TOLAOTMRSELC 165

RESULT 13

O9ZDKO PRELIMINARY: PRT: 311 AA.

DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE HYPOTHETICAL 34.0 KDA PROTEIN.
 GN RP328.
 OS Rickettsia prowazekii.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=782;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAURID E;
 RA MEDLINE=99039499; PubMed=9823893;
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
 RA Sichteritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
 RT "The genome sequence of Rickettsia prowazekii and the origin of mitochondria."
 RL Nature 396:133-140(1998).
 DR EMBL: AJ235271; CAI4788.1; -
 DR InterPro: IPR001107; Band_7.
 DR InterPro: IPR001972; Stomatin.
 DR Pfam: PF01145; Band_7; 1.
 DR PRINTS: PR00721; STOMATIN.
 DR SMART: SM00244; PHB; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 311 AA; 34040 MW; F74F625EC567E5CA CRC64;

Query Match 3.1%; Score 11; DB 16; Length 311;
 Best Local Similarity 100.0%; Pred. No. 0.028;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 127 TOLAOTMRSE 137
 DB 109 TOLAOTMRSE 119

RESULT 14

O92IG8 PRELIMINARY: PRT: 312 AA.

DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE HYPOTHETICAL PROTEIN RC0452.
 GN RC0452.
 OS Rickettsia conorii.

OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=781;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MALISH 7;
 RX MEDLINE=21442074; PubMed=11557893;
 RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
 RA Samsom D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
 RA Raoult D.;
 RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii."
 RL Science 293:2093-2098(2001).
 DR EMBL: AE008609; AL102990.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 312 AA; 34025 MW; A1878CANAABC49F5 CRC64;

Query Match 3.1%; Score 11; DB 16; Length 312;
 Best Local Similarity 100.0%; Pred. No. 0.028;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 127 TOLAOTMRSE 137
 DB 109 TOLAOTMRSE 119

RESULT 15

O60121 PRELIMINARY: PRT: 354 AA.

DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE HYPOTHETICAL 39.3 KDA PROTEIN C16G5.07C IN CHROMOSOME II.
 GN SPG16G5.07C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetiales; Schizosaccharomycetaceae;
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Lyne M., Wood V., Rajandream M.A., Barrell B.G., Beck A.,
 RA Reinhardt R.;
 RT Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE BAND 7 / MEC-2 FAMILY.
 DR EMBL: AL023554; CAI19027.1; -
 DR InterPro: IPR001107; Band_7.
 DR InterPro: IPR001972; Stomatin.
 DR Pfam: PF01145; Band_7; 1.
 DR SMART: SM00244; PHB; 1.
 DR PROSITE: PS01270; BAND_7; FALSE_NEG.
 KW Hypothetical protein.
 SQ SEQUENCE 354 AA; 39274 MW; 66B09E6A12BDC030 CRC64;

Query Match 3.1%; Score 11; DB 3; Length 354;
 Best Local Similarity 100.0%; Pred. No. 0.032;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 111 DPKASYGVED 121
 DB 126 DPKASYGVED 136

Search completed: September 22, 2002, 18:36:37
 Job time: 294 sec

RT "prohibitins, stomatins, and plant disease response genes comprise a
RT protein superfamily that controls cell proliferation, ion channel
RT regulation, and death."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF236372; AAF68388.1; -
DR InterPro: IPR001107; Band_7
DR InterPro: IPR001972; Stomatlin.
DR Pfam: PF01145; Band_7; 1.
DR PRINTS: PR00721; STOMATIN.
DR SMART: SM00244; PHB; 1.
SQ SEQUENCE 394 AA: 43277 MW: 52A409DCFI5C5A45 CRC64;

Query Match 3.7%; Score 13; DB 10; Length 394;
Best Local Similarity 100.0%; Pred. No. 0.00037;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 128 QLAQTTMRSELGK 140
DB 148 QLAQTTMRSELGK 160

RESULT 10
O9LW0

PRELIMINARY: PRT; 401 AA.

AC O9LW0: PRELIMINARY: PRT; 401 AA.
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
DE SIMILARITY TO STOMATIN LIKE PROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and YAC
RT clones."
RL DNA Res. 7:31-63(2000).
RL EMBL: AB018115; BAA97132.1; -
DR InterPro: IPR001107; Band_7.
DR InterPro: IPR001972; Stomatlin.
DR Pfam: PF01145; Band_7; 1.
DR PRINTS: PR00721; STOMATIN.
DR SMART: SM00244; PHB; 1.
SQ SEQUENCE 401 AA: 43607 MW: 28E8BD732A370AB6 CRC64;

Query Match 3.7%; Score 13; DB 10; Length 401;
Best Local Similarity 100.0%; Pred. No. 0.00038;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 128 QLAQTTMRSELGK 140
DB 196 QLAQTTMRSELGK 208

RESULT 11
O93VP9

PRELIMINARY: PRT; 411 AA.

AC O93VP9: PRELIMINARY: PRT; 411 AA.
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE HYPOTHETICAL 45.0 KDA PROTEIN.
GN AT4G27580.
OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamuya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shinn P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,
RA Ecker J.R., Theologis A.;
RT "Full length cDNA of gene AT4G27580 (GI:7269612).";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

[2]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,
RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,
RA Jones T., Kamuya A., Karlin-Neumann G., Kawai J., Kim C., Koesema E.,
RA Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,
RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "Full length cDNA of gene AT4G27580 (GI:7269612).";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY059109; AAL15215.1; -
DR EMBL: AY034924; AAK59431.1; -
KW Hypothetical protein.
SQ SEQUENCE 411 AA: 45020 MW: 94E24ED4C9B89633 CRC64;

Query Match 3.7%; Score 13; DB 10; Length 411;
Best Local Similarity 100.0%; Pred. No. 0.00038;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 128 QLAQTTMRSELGK 140
DB 153 QLAQTTMRSELGK 165

RESULT 12
O9T082

PRELIMINARY: PRT; 515 AA.

ID O9T082: PRELIMINARY: PRT; 515 AA.
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
DE HYPOTHETICAL 55.9 KDA PROTEIN.
GN T29A15.70 OR AT4G27580.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Van Der Schueren J., Chuang Y.-J., Voel M., Robben J.,
RA Volckaert G., Hohnleis J., Mewes H.W., Mayer K.F.X., Schueller C.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Robben J., Grymoprez B., Volckaert G., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;

[illegible]

```

DR PRINTS; PR00721; STOMATIN.
DR SMART; SMO0244; PHB; 1.
SQ SEQUENCE 323 AA; 35668 MW; D5C6241445FEF4DE CRC64;

Query Match          5.6%; Score 20; DB 5; Length 323;
Best Local Similarity 100.0%; Pred. No. 3.8e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 40 FVPOGEAWVERMGGRHRL 59
      |||||||
DB 2 FVPOGEAWVERMGGRHRL 21

RESULT 8
O9XVP9 PRELIMINARY: PRT: 334 AA.
ID O9XVP9
AC O9XVP9;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE F30A10.5 PROTEIN.
GN F30A10.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodertinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Barlow K.;
RL Submitted (CCr-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
   investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z81072; CAB03018.1;
DR InterPro; IPR001107; Band_7;
DR InterPro; IPR001972; Stomatlin.
DR Pfam; PF01145; Band_7; 1.
DR PRINTS; PR00721; STOMATIN.
DR SMART; SMO0244; PHB; 1.
SQ SEQUENCE 334 AA; 36704 MW; 10F98BB9AB44E5ED CRC64;

Query Match          3.9%; Score 14; DB 5; Length 334;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0

OY 40 FVPOGEAWVERMG 53
      |||||||
DB 47 FVPOGEAWVERMG 60

RESULT 9
O9M585 PRELIMINARY: PRT: 394 AA.
ID O9M585
AC O9M585;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE STOMATIN-LIKE PROTEIN.
GN STM1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Nadimpalli R., Yalpani N., Jhal G.S., Simmons C.R.;
```

```

RESULT 4
ID 09P042 PRELIMINARY: PRT: 342 AA.
AC 09P042:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HSP108.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
NCBI_TaxID=9606;
RN 11)
RP SEQUENCE FROM N.A.
RC TISSUE=BLD;
RX MEDLINE-2049367; PubMed-11042152;
RA Zhang Q.H., Ye M., Wu X.Y., Ren S.X., Zhao M., Zhao C.J., Fu G.,
Shen Y., Fan H.Y., Lu G., Zhong M., Xu X.R., Han Z.G., Zhang J.W.,
Tao J., Huang Q.H., Zhou J., Hu G.X., Gu J., Chen S.J., Chen Z.;
RT "Cloning and functional analysis of cDNAs with open reading frames for
RT 300 previously undefined genes expressed in CD34+ hematopoietic
RT stem/progenitor cells."
RL Genome Res. 10:1546-1560(2000).
DR EMBL: AF161456; AAF29073.1; -.
DR InterPro: IPR001107; Band_7.
DR InterPro: IPR001972; Stomatlin.
DR Pfam: PF01145; Band_7; 1.
DR PRINTS: PR00721; STOMATLIN.
DR SMART: SM00244; PHB: 1.
SO SEQUENCE 342 AA; 37145 MW; 79580E3BDBE53E5 CRC64;

Query Match 40.2%; Score 143; DB 4; Length 342;
Best Local Similarity 100.0%; Pred. No. 1.7e-133;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 203 TVLESGTRESAINVAEGKKAQOILASEAKAEQINQAGEASAVLAKAKAKAEAIRILA 262
DB 189 TVLESGTRESAINVAEGKKAQOILASEAKAEQINQAGEASAVLAKAKAKAEAIRILA 248
QY 263 AALTQNGDAASLTVAEQYVSFSLKADSNITLLPSPNGDVTSMVAQAMGYGALTGA 322
DB 249 AALTQNGDAASLTVAEQYVSFSLKADSNITLLPSPNGDVTSMVAQAMGYGALTGA 308
QY 323 PVPGTDSLSSGSSRDVQGTDA 345
DB 309 PVPGTDSLSSGSSRDVQGTDA 331

RESULT 5
ID 09DCG8 PRELIMINARY: PRT: 353 AA.
AC 09DCG8:
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE 0610038F01RIK PROTEIN.
GN 0610038F01RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN 11)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=KIDNEY;
RX MEDLINE-21083660; PubMed-11217851;
RA Kawai J., Shitagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

```

```

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nixaldo I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Bash G.,
RA Lake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamlya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch R.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Williams L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL: AK002793; BAB22363.1; -.
DR MGD: MGI:1913842; 0610038F01RIK.
DR InterPro: IPR001107; Band_7.
DR InterPro: IPR001972; Stomatlin.
DR Pfam: PF01145; Band_7; 1.
DR PRINTS: PR00721; STOMATLIN.
DR SMART: SM00244; PHB: 1.
SO SEQUENCE 353 AA; 38530 MW; D317CBB7E32F8863 CRC64;

Query Match 29.8%; Score 106; DB 11; Length 353;
Best Local Similarity 100.0%; Pred. No. 5.6e-96;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 IYDAINQADCMGICRLREIKIHHPYRKESMOMQVEAEKRRKATVLESEGTRESAIN 216
DB 157 IYDAINQADCMGICRLREIKIHHPYRKESMOMQVEAEKRRKATVLESEGTRESAIN 216
QY 217 VAEGRKQAOILASEAKAEQINQAGEASAVLAKAKAKAEAIRILA 262
DB 217 VAEGRKQAOILASEAKAEQINQAGEASAVLAKAKAKAEAIRILA 262

RESULT 6
ID 099JB2 PRELIMINARY: PRT: 353 AA.
AC 099JB2:
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE RIKEN CDNA 0610038F01 GENE (STOMATLIN-LIKE PROTEIN 2).
GN 0610038F01RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN 11)
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RC STRAIN=FVB; TISSUE=KIDNEY;
RP SEQUENCE FROM N.A.
RA Chang J.G., Chan W.L.;
RT "Mouse stomatin-like protein 2 (MSLP2) mRNA."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC003425; AAH03425.1; -.
DR EMBL: AF323178; AAG53404.1; -.
DR MGD: MGI:1913842; 0610038F01RIK.
DR InterPro: IPR001107; Band_7.
DR InterPro: IPR001972; Stomatlin.
DR Pfam: PF01145; Band_7; 1.
DR PRINTS: PR00721; STOMATLIN.
DR SMART: SM00244; PHB: 1.
SO SEQUENCE 353 AA; 38385 MW; 391D269576F6E6BB CRC64;

Query Match 29.8%; Score 106; DB 11; Length 353;

```

RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahata K., Masuho Y.,
 RA Niimura K., Iwayanagi T.,
 RT "NEDO human cDNA sequencing project."
 RT Submitted (May-2001) to the EMBL/Genbank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PANCREAS, EPITHELIOID CARCINOMA;
 RA Strausberg R.;
 RA Submitted (Oct-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF190167; AAF09142.1; -
 DR EMBL: AF282596; AAF91466.1; -
 DR EMBL: BC003025; AAH03025.1; -
 DR EMBL: BC002442; AAH02442.1; -
 DR EMBL: AK027405; BAB55091.1; -
 DR EMBL: BC014990; AAH14990.1; -
 DR InterPro: IPR001107; Band_7.
 DR InterPro: IPR001972; Stomatin.
 DR Pfam: PF01145; Band_7; 1.
 DR PRINTS: PR00721; STOMATIN.
 DR SMART: SM00244; PHB; 1.
 SQ SEQUENCE 356 AA; 38534 MW; 672331B57C82654E CRC64;

Query Match 61.2%; Score 218; DB 4; Length 356;
 Best Local Similarity 99.7%; Pred. No. 1.7e-206;
 Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 27 RASSGLPRTVVLFPVPOEAMVVERKGRFRILFPGNLILFVLRIRYQSLKEIVNV 86
 |||||
 DB 27 RASSGLPRTVVLFPVPOEAMVVERKGRFRILFPGNLILFVLRIRYQSLKEIVNV 86
 |||||

QY 87 PQSAVTLNDVTLQIDGVLYLRIMPYKASVGEDEPEYAVTQLAOTTMSELSXDKY 146
 |||||
 DB 87 PQSAVTLNDVTLQIDGVLYLRIMPYKASVGEDEPEYAVTQLAOTTMSELSXDKY 146
 |||||

QY 147 FFERESLNASYDAINQADACGIRCLREIKDIHVPVVKESMOMQVAEERKRATVLE 206
 |||||
 DB 147 FFERESLNASYDAINQADACGIRCLREIKDIHVPVVKESMOMQVAEERKRATVLE 206
 |||||

QY 207 SEGTRESAIVNAEGKKQAQIILASEAKAQIINQAGEASAVLAKAKAEALRIILAA 266
 |||||
 DB 207 SEGTRESAIVNAEGKKQAQIILASEAKAQIINQAGEASAVLAKAKAEALRIILAA 266
 |||||

QY 267 QHNGDAASLTVAEOYVSFAFSKLAKDSNTILLPSNPGDVTSMVAQAMGYGALTRAPVG 326
 |||||
 DB 267 QHNGDAASLTVAEOYVSFAFSKLAKDSNTILLPSNPGDVTSMVAQAMGYGALTRAPVG 326
 |||||

QY 327 TPDSLSGSSRDVQGTAS 345
 |||||
 DB 327 TPDSLSGSSRDVQGTAS 345
 |||||

RESULT 2
 Q96FY2 PRELIMINARY; PRT; 356 AA.
 AC Q96FY2;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DE STOMATIN-LIKE 2.
 OS Homo sapiens (Human).
 OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Eukaryota; Eutera; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LUNG CARCINOMA;
 RA Strausberg R.;
 RT Submitted (Jul-2001) to the EMBL/Genbank/DBJ databases.

DR EMBL: BC010152; AAH10152.1; -
 SQ SEQUENCE 356 AA; 38518 MW; 6FC0BF899E1919A7 CRC64;
 Query Match 56.7%; Score 202; DB 4; Length 356;
 Best Local Similarity 100.0%; Pred. No. 1e-190;
 Matches 202; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 DKVFERESLNASYDAINQADACGIRCLREIKDIHVPVVKESMOMQVAEERKRAT 203
 |||||
 DB 144 DKVFERESLNASYDAINQADACGIRCLREIKDIHVPVVKESMOMQVAEERKRAT 203
 |||||

QY 204 VLESEGTRESAIVNAEGKKQAQIILASEAKAQIINQAGEASAVLAKAKAEALRIILAA 263
 |||||
 DB 204 VLESEGTRESAIVNAEGKKQAQIILASEAKAQIINQAGEASAVLAKAKAEALRIILAA 263
 |||||

QY 264 ALTOHNGDAASLTVAEOYVSFAFSKLAKDSNTILLPSNPGDVTSMVAQAMGYGALTRAP 323
 |||||
 DB 264 ALTOHNGDAASLTVAEOYVSFAFSKLAKDSNTILLPSNPGDVTSMVAQAMGYGALTRAP 323
 |||||

QY 324 VPGETPSLSGSSRDVQGTAS 345
 |||||
 DB 324 VPGETPSLSGSSRDVQGTAS 345
 |||||

RESULT 3
 O60376 PRELIMINARY; PRT; 357 AA.
 AC O60376;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PL11659.4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutera; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=FIBROBLAST OR FORESKIN;
 RA Lamerdin J.E., McCreedy P.M., Skowronski E., Adamson A.W.,
 RA Burthart-Schultz K., Gordon L., Kyle A., Ramirez M., Stillwagen S.,
 RA Phan H., Velasco N., Ganes J., Dangnan L., Poundstone P.,
 RA Christensen M., Georges A., Avila J., Liu S., Attix C., Andreise T.,
 RA Trankheim M., Amico-Keller G., Coeffield J., Duarte S., Lucas S.,
 RA Bruce R., Thomas P., Quan G., Kronmiller B., Arellano A.,
 RA Montgomery M., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.O.,
 RA Carrano A.V.;
 RT "Sequence analysis of a human P1 clone containing the XRC9 DNA repair
 RT gene."
 RT Submitted (Mar-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AC004472; AAC07983.1; -
 DR InterPro: IPR001107; Band_7.
 DR Pfam: PF01145; Band_7; 1.
 DR SMART: SM00244; PHB; 1.
 SQ SEQUENCE 357 AA; 38749 MW; 512632B83028135A CRC64;

Query Match 42.7%; Score 152; DB 4; Length 357;
 Best Local Similarity 100.0%; Pred. No. 2.3e-141;
 Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 VEAERKRATVLESEGTRESAIVNAEGKKQAQIILASEAKAQIINQAGEASAVLAKAKA 253
 |||||
 DB 194 VEAERKRATVLESEGTRESAIVNAEGKKQAQIILASEAKAQIINQAGEASAVLAKAKA 254
 |||||

QY 254 KAEALRIILAAALTOHNGDAASLTVAEOYVSFAFSKLAKDSNTILLPSNPGDVTSMVAQAM 313
 |||||
 DB 254 KAEALRIILAAALTOHNGDAASLTVAEOYVSFAFSKLAKDSNTILLPSNPGDVTSMVAQAM 314
 |||||

QY 314 GYGALTRAPVGETPSLSGSSRDVQGTAS 345
 |||||
 DB 314 GYGALTRAPVGETPSLSGSSRDVQGTAS 346
 |||||

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 18:31:43 : Search time 25.31 Seconds
(without alignments)
2433.275 Million cell updates/sec

Title: US-09-898-216-1

Perfect score: 356
Sequence: 1 MLARARGHMGPFAGLSTG.....RDVGTDPASXDELDRVKMS 356

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 17294929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	218	61.2	356	4	Q9UJZ1
2	202	56.7	356	4	Q96FY2
3	152	42.7	357	4	O60376
4	143	40.2	342	4	O9P042
5	106	29.8	353	11	Q9DC68
6	106	29.8	353	11	Q99J82
7	20	5.6	323	5	Q9W1F7
8	14	3.9	334	5	Q9XVP9
9	13	3.7	394	10	Q9M585
10	13	3.7	401	10	Q9LWV0
11	13	3.7	411	10	Q93VP9
12	13	3.7	515	10	Q9T082
13	11	3.1	511	16	Q9ZDK0
14	11	3.1	312	16	Q92IC8
15	11	3.1	354	3	O60121
16	9	2.5	357	5	Q9U4M5

17	9	2.5	374	5	077376	077376 plasmodium
18	8	2.2	235	9	Q94M69	Q94M69 streptococc
19	8	2.2	248	16	Q9ABD4	Q9ABD4 caulobacter
20	8	2.2	263	16	Q9HXE1	Q9HXE1 pseudomonas
21	8	2.2	398	10	O04471	O04471 arabidopsis
22	8	2.2	477	5	O62022	O62022 caenorhabdi
23	8	2.2	523	5	O77238	O77238 drosophila
24	8	2.2	554	5	Q9V7X7	Q9V7X7 drosophila
25	8	2.2	658	5	O18703	O18703 caenorhabdi
26	8	2.2	716	9	O80256	O80256 vibrio chol
27	8	2.2	754	5	Q9NED3	Q9NED3 leishmania
28	8	2.2	1443	5	Q9VPR0	Q9VPR0 drosophila
29	8	2.2	8817	2	O53840	O53840 polyanthum
30	7	2.0	47	6	Q28882	Q28882 ovis aries
31	7	2.0	95	5	Q26830	Q26830 trypanosoma
32	7	2.0	95	16	Q91054	Q91054 pseudomonas
33	7	2.0	110	2	Q9APJ7	Q9APJ7 hyphomicrob
34	7	2.0	112	16	O26072	O26072 helicobacte
35	7	2.0	112	16	Q92J64	Q92J64 helicobacte
36	7	2.0	120	5	Q26834	Q26834 trypanosoma
37	7	2.0	121	17	Q9YC53	Q9YC53 aeropyrum p
38	7	2.0	129	16	Q98AP9	Q98AP9 rhizobium l
39	7	2.0	133	2	Q939R0	Q939R0 fibrobacter
40	7	2.0	134	11	O88728	O88728 mus musculu
41	7	2.0	135	5	Q9TXE2	Q9TXE2 chlamys nlp
42	7	2.0	135	17	Q9YBJ8	Q9YBJ8 aeropyrum p
43	7	2.0	136	16	Q92MW3	Q92MW3 rhizobium m
44	7	2.0	137	15	Q91HW7	Q91HW7 human immun
45	7	2.0	142	11	Q9CSJ6	Q9CSJ6 mus musculu

ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	356 AA.
1	Q9UJZ1			
AC	Q9UJZ1;			
DT	01-MAY-2000 (Tremblrel, 13, Created)			
DT	01-MAY-2000 (Tremblrel, 13, last sequence update)			
DT	01-DEC-2001 (Tremblrel, 19, last annotation update)			
DE	MEMBRANE ASSOCIATED PROTEIN SLP-2 (STOMATIN-LIKE PROTEIN 2) (STOMATIN-LIKE 2) (CDNA FLJ14499 FIS, CLONE NTZRM100080, WEAKLY SIMILAR TO UNC-1 PROTEIN) (STOMATIN-LIKE 2).			
DE	UNC-1 PROTEIN) (STOMATIN-LIKE 2).			
GN	HUSLP2 OR SLP2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RP	(1)			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=HEART MUSCLE;			
RA	Wang Y., Morrow J.S.;			
RT	"Identification and Characterization of Human SLP-2, a Novel Homologue of Stomatlin (band 7.2b) Present in Erythrocytes and Other Tissues.";			
RL	Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.			
RN	(2)			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=BRN;			
RA	Owczarek C.M., Treutlein H.R., Portbury K.J., Gulluyan L.M., Kola I., Hertzog P.J.;			
RT	"A novel member of the stomatin/EPB72/MEC-2 family, stomatin-like protein 2 (SLP2), is ubiquitously expressed and localizes to HSA chromosome 9p13.1.";			
RT	Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.			
RN	(3)			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=LUNG CARCINOMA, SKIN, AND MELANOTIC MELANOMA.;			
RA	Strausberg R.;			
RL	Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.			
RN	(4)			
RP	SEQUENCE FROM N.A.			
RA	Isoagai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,			

QY 182 -----VPPRVKESM----- 190
DB 132 IGVAEABDAGIRAEACCKEMLDVKMDPTKIADSKRAPELOKSAFSEVNITAEAOQA 191
QY 191 -----OMOVAERKRATYLE-----SEGTRESAINV 217
DB 192 YELGAREOQKROEIEIEVQKKOIAVEAOEILRTDKELIATVRRPAEAEHRIQOI 251
QY 218 AEGKKOAOILASAEKAEQINOAGASAVLAKAKAEAIRI-LAAALTCNHGDAASL 276
DB 252 ABECKVQYLLAOAE-AEKIRK-IGEAEAIVLEAMGKAERKKLAEAVQKYGDAAKMA 309
QY 277 TWAEOYVSASFSLA-----KDSNTILLPSNPGDVTSMVAQMGVYGALTKAPVGPDPDLS 332
DB 310 LVLEALPOIAAKIAAPLTVDELIVLSGDSKVTSEVNRLL-----AELPASVHALT 361
QY 333 SGSSRDVQGTDS 345
DB 362 -----GVDLS 366

RESULT 8
US-08-762-106-8
; Sequence 8, Application US/08762106
; Patent No. 5948677
; GENERAL INFORMATION:
; APPLICANT: Jarvik, Jonathan W.
; TITLE OF INVENTION: READING FRAME INDEPENDENT EPIOTOPE
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harris Brotman
; STREET: 202 Coast Blvd., Suite 111
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/762,106
; FILING DATE: 09-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brotman, Harris F.
; REGISTRATION NUMBER: 35,461
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 654-2428
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 516 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: Internal
; US-08-762-106-8

Query Match 5.9%; Score 104.5; DB 2; Length 516;
Best Local Similarity 25.6%; Pred. No. 0.067;
Matches 74; Conservative 33; Mismatches 95; Indels 87; Gaps 15;
QY 55 FHRILEPGLNLLIPVLDRI--YVOSIKEIVINVPQSAVTLQNDVTLQIDGVLRLIMP 112
DB 257 FDLTVSSGY-IYDPVMEVETAFMPMLKEQAIQYLAR-----GVVARRVVDK 302
QY 113 YKASYGVEDEPVAVTQLAQTMRSELGKLSXDKVFRRESLNASIYDAINQADQCWIRC 172
DB 303 L-----VED-----AAALAN--RSTLA-----DKAATAATYDA-----WAERQ 336

QY 173 LRYEIKDHPVPRVKESMQOVEERRKRATYLE-----SEGTRESAINVAEGKQAO 225
DB 337 AKME-----ALQCKELEAVRRRPTFVLRBLKPAVASADAVPA--AAELTQAE 364
QY 226 ILAS---EAEKAOINOAGASAVLAKAKAEAIRITLAAALTOHNGDAASLTVAEOY 282
DB 385 EAANNAKWEADKAEEAKAEKAAEAAAEQKALLEELATATAAEAEERGEPPPAEPPS--- 441
QY 283 VSAFSKLAKNSNTLLPS--NPGDVTSMVAQMGVYGALTKAPVGPDPD 329
DB 442 -----LPDGVPEVDVEAEVAKAV---EAVKPPVKEVTD 472

RESULT 9
US-08-745-404-2
; Sequence 2, Application US/08745404B
; Patent No. 6096717
; GENERAL INFORMATION:
; APPLICANT: Jarvik, Jonathan W.
; TITLE OF INVENTION: Method for Producing Tagged Genes,
; FILE REFERENCE: 2087-961422
; CURRENT APPLICATION NUMBER: US/08/745,404B
; EARLIER FILING DATE: 1996-11-08
; EARLIER APPLICATION NUMBER: 08/000,619
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 516
; TYPE: PRP
; ORGANISM: Chlamydomonas
; US-08-745-404-2

Query Match 5.9%; Score 104.5; DB 3; Length 516;
Best Local Similarity 25.6%; Pred. No. 0.067;
Matches 74; Conservative 33; Mismatches 95; Indels 87; Gaps 15;
QY 55 FHRILEPGLNLLIPVLDRI--YVOSIKEIVINVPQSAVTLQNDVTLQIDGVLRLIMP 112
DB 257 FDLTVSSGY-IYDPVMEVETAFMPMLKEQAIQYLAR-----GVVARRVVDK 302
QY 113 YKASYGVEDEPVAVTQLAQTMRSELGKLSXDKVFRRESLNASIYDAINQADQCWIRC 172
DB 303 L-----VED-----AAALAN--RSTLA-----DKAATAATYDA-----WAERQ 336
QY 173 LRYEIKDHPVPRVKESMQOVEERRKRATYLE-----SEGTRESAINVAEGKQAO 225
DB 337 AKME-----ALQCKELEAVRRRPTFVLRBLKPAVASADAVPA--AAELTQAE 364
QY 226 ILAS---EAEKAOINOAGASAVLAKAKAEAIRITLAAALTOHNGDAASLTVAEOY 282
DB 385 EAANNAKWEADKAEEAKAEKAAEAAAEQKALLEELATATAAEAEERGEPPPAEPPS--- 441
QY 283 VSAFSKLAKNSNTLLPS--NPGDVTSMVAQMGVYGALTKAPVGPDPD 329
DB 442 -----LPDGVPEVDVEAEVAKAV---EAVKPPVKEVTD 472

RESULT 10
US-09-320-774-8
; Sequence 8, Application US/09320774
; Patent No. 6265545
; GENERAL INFORMATION:
; APPLICANT: Jarvik, Jonathan W.
; TITLE OF INVENTION: READING FRAME INDEPENDENT EPIOTOPE
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harris Brotman
; STREET: 202 Coast Blvd., Suite 111

CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/320,774
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/762,106
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brotman, Harris F.
REGISTRATION NUMBER: 35,461
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 654-2428
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 516 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-09-320-774-8

Query Match 5.9%; Score 104.5; DB 4; Length 516;
Best Local Similarity 25.6%; Pred. No. 0.067;
Matches 74; Conservative 33; Mismatches 95; Indels 87; Gaps 15;

QY 55 FHRIIEPGNIIIPVDRIR--YVQSLKEIYINVPQSAVTLDNVTLOIDGVLYLRIMP 112
DB 255 FRLIVSSG-IYDPVAREVETAFMPWLKQDAIGYLAR-----GVYARRVYVK 302
QY 113 YKASYGVDEPEYAVTQLOATMRSELGKLSXDKVFERESLNASTYDAINQAADCWGIRC 172
DB 303 L-----VED-----AAALAN--RSTLA-----DKAASATAATVDA-----WAERQ 336
QY 173 LYEIKDHIHVPKVESMOMQYEAERKRATYLE-----SEGRTESINVAEGKKQO 225
DB 337 AKME-----AELQKLEAVARRPTFVRLKPAVASDAVEAA--AAELTQAQ 384
QY 226 ILAS---EAEKAEQIQAAGEASAVLAKAKAEAIRIILAAALTOHNGDAASLTVAEOY 282
DB 385 EAAAKWEADKAEAEKAEAEAAAEOKALLEELAAATAAEAEERGEPPAEPPS--- 441
QY 283 VSAFSKLAKDSNTILPS--NPGDVTSMVAQMGVYGALTAKAPVPGTPD 329
DB 442 -----LPDGEPEVDVEAEVAKAV---EAVPKPPVKVETD 472

RESULT 11
US-08-762-106-9
Sequence 9, Application US/08762106
Patent No. 5948677
GENERAL INFORMATION:
APPLICANT: Jarvik, Jonathan W.
TITLE OF INVENTION: READING FRAME INDEPENDENT EPIOTOPE
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Harris Brotman
STREET: 202 Coast Blvd., Suite 111
CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/762,106
FILING DATE: 09-DEC-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brotman, Harris F.
REGISTRATION NUMBER: 35,461
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 654-2428
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-762-106-9

Query Match 5.9%; Score 104.5; DB 2; Length 527;
Best Local Similarity 25.6%; Pred. No. 0.069;
Matches 74; Conservative 33; Mismatches 95; Indels 87; Gaps 15;

QY 55 FHRIIEPGNIIIPVDRIR--YVQSLKEIYINVPQSAVTLDNVTLOIDGVLYLRIMP 112
DB 268 FRLIVSSG-IYDPVAREVETAFMPWLKQDAIGYLAR-----GVYARRVYVK 313
QY 113 YKASYGVDEPEYAVTQLOATMRSELGKLSXDKVFERESLNASTYDAINQAADCWGIRC 172
DB 314 L-----VED-----AAALAN--RSTLA-----DKAASATAATVDA-----WAERQ 347
QY 173 LYEIKDHIHVPKVESMOMQYEAERKRATYLE-----SEGRTESINVAEGKKQO 225
DB 348 AKME-----AELQKLEAVARRPTFVRLKPAVASDAVEAA--AAELTQAQ 395
QY 226 ILAS---EAEKAEQIQAAGEASAVLAKAKAEAIRIILAAALTOHNGDAASLTVAEOY 282
DB 396 EAAAKWEADKAEAEKAEAEAAAEOKALLEELAAATAAEAEERGEPPAEPPS--- 452
QY 283 VSAFSKLAKDSNTILPS--NPGDVTSMVAQMGVYGALTAKAPVPGTPD 329
DB 453 -----LPDGEPEVDVEAEVAKAV---EAVPKPPVKVETD 483

RESULT 12
US-09-320-774-9
Sequence 9, Application US/09320774
Patent No. 6265545
GENERAL INFORMATION:
APPLICANT: Jarvik, Jonathan W.
TITLE OF INVENTION: READING FRAME INDEPENDENT EPIOTOPE
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Harris Brotman
STREET: 202 Coast Blvd., Suite 111
CITY: La Jolla
STATE: California
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/320,774
FILING DATE:
CLASSIFICATION:

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2002, 18:14:17 ; Search time 196.92 Seconds
(without alignments)
636.325 Million cell updates/sec

Title: US-09-898-216-1
Perfect score: 1767
Sequence: 1 MLARARGHWGPFAGLSTG.....RDVQGTDA SXDELDREVKMS 356

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3502263 seqs, 351980561 residues
Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents_AA_Main:*

1:	/cgn2_6/ptodata/1/paa/PCTUS.COMB.pep.*
2:	/cgn2_6/ptodata/1/paa/US06.COMB.pep.*
3:	/cgn2_6/ptodata/1/paa/US07.COMB.pep.*
4:	/cgn2_6/ptodata/1/paa/US08.COMB.pep.*
5:	/cgn2_6/ptodata/1/paa/US082.COMB.pep.*
6:	/cgn2_6/ptodata/1/paa/US083.COMB.pep.*
7:	/cgn2_6/ptodata/1/paa/US084.COMB.pep.*
8:	/cgn2_6/ptodata/1/paa/US085.COMB.pep.*
9:	/cgn2_6/ptodata/1/paa/US086.COMB.pep.*
10:	/cgn2_6/ptodata/1/paa/US087.COMB.pep.*
11:	/cgn2_6/ptodata/1/paa/US088.COMB.pep.*
12:	/cgn2_6/ptodata/1/paa/US089.COMB.pep.*
13:	/cgn2_6/ptodata/1/paa/US090.COMB.pep.*
14:	/cgn2_6/ptodata/1/paa/US091.COMB.pep.*
15:	/cgn2_6/ptodata/1/paa/US092.COMB.pep.*
16:	/cgn2_6/ptodata/1/paa/US093.COMB.pep.*
17:	/cgn2_6/ptodata/1/paa/US094.COMB.pep.*
18:	/cgn2_6/ptodata/1/paa/US095.COMB.pep.*
19:	/cgn2_6/ptodata/1/paa/US096.COMB.pep.*
20:	/cgn2_6/ptodata/1/paa/US097.COMB.pep.*
21:	/cgn2_6/ptodata/1/paa/US098.COMB.pep.*
22:	/cgn2_6/ptodata/1/paa/US099.COMB.pep.*
23:	/cgn2_6/ptodata/1/paa/US100.COMB.pep.*
24:	/cgn2_6/ptodata/1/paa/US101.COMB.pep.*
25:	/cgn2_6/ptodata/1/paa/US60.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1646	93.2	356	1 PCT-US98-12125-145	Sequence 145, App
2	1646	93.2	356	1 US-09-209-462B-150	Sequence 150, App
3	1646	93.2	378	1 PCT-US01-03800A-2351	Sequence 2351, App
4	1646	93.2	378	1 PCT-US01-04098A-3196	Sequence 3196, App
5	1643	93.0	3007	1 PCT-US01-04842A-222	Sequence 222, App
6	1643	93.0	3019	1 PCT-US01-08656-10594	Sequence 10594, App
7	1635	92.5	358	1 PCT-US01-04098A-1228	Sequence 1228, App

8	1493	84.5	306	1 PCT-US98-12125-272	Sequence 272, App
9	1493	84.5	306	1 US-09-209-462B-411	Sequence 411, App
10	1488	84.2	305	16 US-09-298-733-36	Sequence 36, App
11	1488	84.2	305	16 US-09-298-733A-36	Sequence 36, App
12	1488	84.2	305	21 US-09-723-594-36	Sequence 36, App
13	1488	84.2	305	21 US-09-724-497-36	Sequence 36, App
14	1064.5	60.2	323	20 US-09-614-150-5415	Sequence 5415, App
15	1064.5	60.2	323	26 US-60-191-637-5433	Sequence 5433, App
16	946	53.5	463	1 PCT-US01-04942A-598	Sequence 598, App
17	894.5	50.6	753	26 US-60-167-717-5529	Sequence 5529, App
18	826	46.7	411	20 US-09-620-394B-4030	Sequence 4030, App
19	826	46.7	513	21 US-09-708-427-12542	Sequence 12542, App
20	803	45.4	405	21 US-09-708-427-56695	Sequence 56695, App
21	799	45.2	356	21 US-09-708-427-56696	Sequence 56696, App
22	796	45.0	344	21 US-09-708-427-56697	Sequence 56697, App
23	794.5	45.0	394	17 US-09-395-197-2	Sequence 2, App
24	794.5	45.0	394	21 US-09-767-129-2	Sequence 2, App
25	794.5	45.0	436	21 US-09-708-427-57092	Sequence 57092, App
26	791.5	44.8	349	21 US-09-708-427-57093	Sequence 57093, App
27	716	40.5	315	22 US-09-806-866-13	Sequence 13, App
28	716	40.5	315	22 US-09-806-866-15	Sequence 15, App
29	712.5	40.3	315	22 US-09-806-866-17	Sequence 17, App
30	711.5	40.3	316	22 US-09-806-866-11	Sequence 11, App
31	661.5	37.4	316	22 US-09-806-866-11	Sequence 11, App
32	619.5	35.1	275	26 US-60-324-109-32115	Sequence 32115, App
33	588	33.3	283	18 US-09-417-507-32569	Sequence 32569, App
34	584.5	33.1	265	21 US-09-708-427-57094	Sequence 57094, App
35	534.5	30.2	253	20 US-09-620-394B-4031	Sequence 4031, App
36	534.5	30.2	357	21 US-09-708-427-12543	Sequence 12543, App
37	480.5	27.2	279	16 US-09-248-796-16859	Sequence 16859, App
38	480.5	27.2	279	26 US-60-096-409-16859	Sequence 16859, App
39	477.5	27.0	320	19 US-09-543-681A-6978	Sequence 6978, App
40	474	26.8	18	18 US-09-489-039A-13396	Sequence 13396, App
41	471.5	26.7	286	22 US-09-897-516-7262	Sequence 7262, App
42	471.5	26.7	286	26 US-60-215-161-7262	Sequence 7262, App
43	465	26.3	349	21 US-09-739-449-11515	Sequence 11515, App
44	465	26.3	349	22 US-09-803-110-11515	Sequence 11515, App
45	460	26.0	368	23 US-09-902-540-11334	Sequence 11334, App

ALIGNMENTS

RESULT: 1
PCT-US98-12125-145
Sequence 145, Application PC/TUS9812125
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 86 Human Secreted Proteins
NUMBER OF SEQUENCES: 318
CORRESPONDENCE ADDRESS:
ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US98/12125
FILING DATE: Jan 01, 1990
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brooks
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: P2008complete

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 145:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 356 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; PCT-US98-12125-145

Query Match          93.2%; Score 1646; DB 1; Length 356;
Best Local Similarity 95.3%; Pred. No. 3.7e-143;
Matches 341; Conservative 2; Mismatches 11; Indels 4; Gaps 2;

QY 1 MLARAARGHMGPFEEG--LSTGFMPRSGRASSGLPRNTVVLFPQOEAMVVERMGFRHRI 58
   1 MLARAARGTGALLRGLSLASGRAPR--RASSGLPRNTVVLFPQOEAMVVERMGFRHRI 58
Db 1 MLARAARGTGALLRGLSLASGRAPR--RASSGLPRNTVVLFPQOEAMVVERMGFRHRI 58
QY 59 LEPLGNTLIPVLDRIRYVQSLKEIYINVEQSAVTLDNVTLQIDGVLYLRIMDPKASYG 118
   59 LEPLGNTLIPVLDRIRYVQSLKEIYINVEQSAVTLDNVTLQIDGVLYLRIMDPKASYG 118
Db 59 LEPLGNTLIPVLDRIRYVQSLKEIYINVEQSAVTLDNVTLQIDGVLYLRIMDPKASYG 118
QY 119 VEDPEYAVTQLAQTTRSELGKSLDKYFRERESLNASTVDAINQAACWGIKRLRYEIK 178
   119 VEDPEYAVTQLAQTTRSELGKSLDKYFRERESLNASTVDAINQAACWGIKRLRYEIK 178
Db 119 VEDPEYAVTQLAQTTRSELGKSLDKYFRERESLNASTVDAINQAACWGIKRLRYEIK 178
QY 179 DIHVPVRKESMOMOVEERRKRATVLESEGTRESAINVAEKKQAQILASAEKAEQIN 238
   179 DIHVPVRKESMOMOVEERRKRATVLESEGTRESAINVAEKKQAQILASAEKAEQIN 238
Db 179 DIHVPVRKESMOMOVEERRKRATVLESEGTRESAINVAEKKQAQILASAEKAEQIN 238
QY 239 QAAGASAVLAKAKAKAEIRILAAALTOHNGDAAASLTVAEQVYSAFSKLAKDSNTILL 298
   239 QAAGASAVLAKAKAKAEIRILAAALTOHNGDAAASLTVAEQVYSAFSKLAKDSNTILL 298
Db 239 QAAGASAVLAKAKAKAEIRILAAALTOHNGDAAASLTVAEQVYSAFSKLAKDSNTILL 298
QY 239 PSNPEDVTSMAVAQANGVYGTALTKAPVPGTPDLSGSSSRDVGCTASXDELDRYKMS 356
   239 PSNPEDVTSMAVAQANGVYGTALTKAPVPGTPDLSGSSSRDVGCTASXDELDRYKMS 356
Db 239 PSNPEDVTSMAVAQANGVYGTALTKAPVPGTPDLSGSSSRDVGCTASXDELDRYKMS 356

RESULT 2
US-09-209-462B-150
; Sequence 150, Application US/09209462B
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; FILE REFERENCE: P2008P1
; CURRENT APPLICATION NUMBER: US/09/209,462B
; CURRENT FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: PCT/US98/12125
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/049,547
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,548
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,549
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,550
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,566
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,606
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,607
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,608
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,609
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,610
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,611
; PRIOR FILING DATE: 1997-06-13
```

```

; PRIOR APPLICATION NUMBER: 60/050,901
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/052,989
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058,665
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,668
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,669
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,750
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,971
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,972
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,975
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/060,834
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,841
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,844
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,865
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,059
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,060
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 737
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 150
; LENGTH: 356
; TYPE: PPT
; ORGANISM: Homo sapiens
; US-09-209-462B-150

Query Match          93.2%; Score 1646; DB 16; Length 356;
Best Local Similarity 95.3%; Pred. No. 3.7e-143;
Matches 341; Conservative 2; Mismatches 11; Indels 4; Gaps 2;

QY 1 MLARAARGHMGPFEEG--LSTGFMPRSGRASSGLPRNTVVLFPQOEAMVVERMGFRHRI 58
   1 MLARAARGTGALLRGLSLASGRAPR--RASSGLPRNTVVLFPQOEAMVVERMGFRHRI 58
Db 1 MLARAARGTGALLRGLSLASGRAPR--RASSGLPRNTVVLFPQOEAMVVERMGFRHRI 58
QY 59 LEPLGNTLIPVLDRIRYVQSLKEIYINVEQSAVTLDNVTLQIDGVLYLRIMDPKASYG 118
   59 LEPLGNTLIPVLDRIRYVQSLKEIYINVEQSAVTLDNVTLQIDGVLYLRIMDPKASYG 118
Db 59 LEPLGNTLIPVLDRIRYVQSLKEIYINVEQSAVTLDNVTLQIDGVLYLRIMDPKASYG 118
QY 119 VEDPEYAVTQLAQTTRSELGKSLDKYFRERESLNASTVDAINQAACWGIKRLRYEIK 178
   119 VEDPEYAVTQLAQTTRSELGKSLDKYFRERESLNASTVDAINQAACWGIKRLRYEIK 178
Db 119 VEDPEYAVTQLAQTTRSELGKSLDKYFRERESLNASTVDAINQAACWGIKRLRYEIK 178
QY 179 DIHVPVRKESMOMOVEERRKRATVLESEGTRESAINVAEKKQAQILASAEKAEQIN 238
   179 DIHVPVRKESMOMOVEERRKRATVLESEGTRESAINVAEKKQAQILASAEKAEQIN 238
Db 179 DIHVPVRKESMOMOVEERRKRATVLESEGTRESAINVAEKKQAQILASAEKAEQIN 238
QY 239 QAAGASAVLAKAKAKAEIRILAAALTOHNGDAAASLTVAEQVYSAFSKLAKDSNTILL 298
   239 QAAGASAVLAKAKAKAEIRILAAALTOHNGDAAASLTVAEQVYSAFSKLAKDSNTILL 298
Db 239 QAAGASAVLAKAKAKAEIRILAAALTOHNGDAAASLTVAEQVYSAFSKLAKDSNTILL 298
QY 239 PSNPEDVTSMAVAQANGVYGTALTKAPVPGTPDLSGSSSRDVGCTASXDELDRYKMS 356
   239 PSNPEDVTSMAVAQANGVYGTALTKAPVPGTPDLSGSSSRDVGCTASXDELDRYKMS 356
Db 239 PSNPEDVTSMAVAQANGVYGTALTKAPVPGTPDLSGSSSRDVGCTASXDELDRYKMS 356

RESULT 3
```


Best Local Similarity 95.0%; Pred. No. 2e-141;
Matches 340; Conservative 3; Mismatches 11; Indels 4; Gaps 2;

```
OY 1 MLARAARGHMGFPAGE--LSTGFWPRSGRASSGLPNTYVLLFVPOQEAANVVERMGFRHRI 58
DB 1 MLARAARGGALLLRSLSLASGRAP--RASSGLPNTYVLLFVPOQEAANVVERMGFRHRI 58
OY 59 LEPGLNILLPVLDRIYVOSLKEIVINPEQSAVTLDNVTLQIDGVLIRINDPYKASYG 118
DB 59 LEPGLNILLPVLDRIYVOSLKEIVINPEQSAVTLDNVTLQIDGVLIRINDPYKASYG 118
OY 119 VEDPEYAVTQLAQTTMRSELGKLSXDKVFRERESLNASIVDAINQAADCGIRCLREYEIK 178
DB 119 VEDPEYAVTQLAQTTMRSELGKLSXDKVFRERESLNASIVDAINQAADCGIRCLREYEIK 178
OY 179 DIHVPRVESMOMQVEARRKRAVLESEGTRESAINVAEGKKQAOILLASEAKEAOIN 238
DB 179 DIHVPRVESMOMQVEARRKRAVLESEGTRESAINVAEGKKQAOILLASEAKEAOIN 238
OY 239 QAAGEASAVLAKAKAKAEAIRILAAALTOHNGDPAASLTVAEOYVSAFSKLAKDSNTILL 298
DB 239 QAAGEASAVLAKAKAKAEAIRILAAALTOHNGDPAASLTVAEOYVSAFSKLAKDSNTILL 298
OY 299 PSNPGDVTSMVAQAMGVYALTAKAPVPGTPDLSLSSGSSRDVOSTDASXDEIDRVKMS 356
DB 299 PSNPGDVTSMVAQAMGVYALTAKAPVPGTPDLSLSSGSSRDVOSTDASXDEIDRVKMT 356
```

RESULT 6

PCT-US01-08656-10594

Sequence 10594, Application PC/TUS0108656

GENERAL INFORMATION:

APPLICANT: Hyseq, Inc.

TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES

FILE REFERENCE: 21272-066

CURRENT APPLICATION NUMBER: PCT/US01/08656

CURRENT FILING DATE: 2001-04-16

PRIOR APPLICATION NUMBER: 09/522,929

PRIOR FILING DATE: 2000-04-18

PRIOR APPLICATION NUMBER: 09/770,160

PRIOR FILING DATE: 2001-01-26

NUMBER OF SEQ ID NOS: 10994

SOFTWARE: Custom

SEQ ID NO 10594

LENGTH: 3019

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: DOMAIN

LOCATION: (115)..(144)

OTHER INFORMATION: Band 7 protein family proteins domain identified by eMATRIX.

OTHER INFORMATION: accession number BL01270C, p-value=6.745e-17, raw score of 16.91

NAME/KEY: DOMAIN

LOCATION: (2426)..(2903)

OTHER INFORMATION: ATPases associated with various cellular act domain

OTHER INFORMATION: Identified by Pfam, accession name AAA, E-value=3.3e-182, Pfam

NAME/KEY: misc_feature

LOCATION: (1)...(3019)

OTHER INFORMATION: xaa - x or * as defined in Table 2

PCT-US01-08656-10594

Query Match 93.0%; Score 1643; DB 1; Length 3019;
Best Local Similarity 95.0%; Pred. No. 2e-141;

Matches 340; Conservative 3; Mismatches 11; Indels 4; Gaps 2;

```
OY 1 MLARAARGHMGFPAGE--LSTGFWPRSGRASSGLPNTYVLLFVPOQEAANVVERMGFRHRI 58
DB 1 MLARAARGGALLLRSLSLASGRAP--RASSGLPNTYVLLFVPOQEAANVVERMGFRHRI 58
OY 59 LEPGLNILLPVLDRIYVOSLKEIVINPEQSAVTLDNVTLQIDGVLIRINDPYKASYG 118
DB 59 LEPGLNILLPVLDRIYVOSLKEIVINPEQSAVTLDNVTLQIDGVLIRINDPYKASYG 118
```

DB 59 LEPGLNILLPVLDRIYVOSLKEIVINPEQSAVTLDNVTLQIDGVLIRINDPYKASYG 118

OY 119 VEDPEYAVTQLAQTTMRSELGKLSXDKVFRERESLNASIVDAINQAADCGIRCLREYEIK 178

DB 119 VEDPEYAVTQLAQTTMRSELGKLSXDKVFRERESLNASIVDAINQAADCGIRCLREYEIK 178

OY 179 DIHVPRVESMOMQVEARRKRAVLESEGTRESAINVAEGKKQAOILLASEAKEAOIN 238

DB 179 DIHVPRVESMOMQVEARRKRAVLESEGTRESAINVAEGKKQAOILLASEAKEAOIN 238

OY 239 QAAGEASAVLAKAKAKAEAIRILAAALTOHNGDPAASLTVAEOYVSAFSKLAKDSNTILL 298

DB 239 QAAGEASAVLAKAKAKAEAIRILAAALTOHNGDPAASLTVAEOYVSAFSKLAKDSNTILL 298

OY 299 PSNPGDVTSMVAQAMGVYALTAKAPVPGTPDLSLSSGSSRDVOSTDASXDEIDRVKMS 356

DB 299 PSNPGDVTSMVAQAMGVYALTAKAPVPGTPDLSLSSGSSRDVOSTDASXDEIDRVKMT 356

RESULT 7

PCT-US01-04098A-1228

Sequence 1228, Application PC/TUS0104098A

GENERAL INFORMATION:

APPLICANT: Hyseq, Inc.

TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides

FILE REFERENCE: 21272-029

CURRENT APPLICATION NUMBER: PCT/US01/04098A

CURRENT FILING DATE: 2001-02-05

PRIOR APPLICATION NUMBER: Not yet Assigned

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: 09/728,422

PRIOR FILING DATE: 2000-11-30

PRIOR APPLICATION NUMBER: 09/693,325

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 09/663,561

PRIOR FILING DATE: 2000-09-15

PRIOR APPLICATION NUMBER: 09/654,936

PRIOR FILING DATE: 2000-09-01

PRIOR APPLICATION NUMBER: 09/620,325

PRIOR FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/598,075

PRIOR FILING DATE: 2000-06-20

PRIOR APPLICATION NUMBER: 09/560,875

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: 09/496,914

PRIOR FILING DATE: 2000-02-03

NUMBER OF SEQ ID NOS: 3960

SOFTWARE: Custom

SEQ ID NO 1228

LENGTH: 358

TYPE: PRT

ORGANISM: Homo sapiens

PCT-US01-04098A-1228

Query Match 92.5%; Score 1635; DB 1; Length 358;
Best Local Similarity 94.7%; Pred. No. 3.9e-142;

Matches 341; Conservative 2; Mismatches 11; Indels 6; Gaps 3;

```
OY 1 MLARAARGHMGFPAGE--LSTGFWPRSGRASSGLPNTYVLLFVPOQEAANVVERMGFRHRI 58
DB 1 MLARAARGGALLLRSLSLASGRAP--RASSGLPNTYVLLFVPOQEAANVVERMGFRHRI 58
OY 59 LEPGLNILLPVLDRIYVOSLKEIVINPEQSAVTLDNVTLQIDGVLIRINDPYKASYG 118
DB 59 LEPGLNILLPVLDRIYVOSLKEIVINPEQSAVTLDNVTLQIDGVLIRINDPYKASYG 118
OY 119 VEDPEYAVTQLAQTTMRSELGKLSXDKVFRERESLNASIVDAINQAADCGIRCLREYEIK 178
DB 119 VEDPEYAVTQLAQTTMRSELGKLSXDKVFRERESLNASIVDAINQAADCGIRCLREYEIK 178
OY 179 DIHVPRVESMOMQVEARRKRAVLESEGTRESAINVAEGKKQAOILLASEAKEAOIN 238
DB 179 DIHVPRVESMOMQVEARRKRAVLESEGTRESAINVAEGKKQAOILLASEAKEAOIN 238
```

Db 179 DIHVPRKESQMOVEAEERRKRAIVLESEGTRESAINVAEKKQAQIILASEAKAEQIN 238
QY 239 QAAGASAVLAKAKAKAEIRILAAALTOHNGDAASLTVAEQYSAFSKILAKDSNTIIL 298
Db 239 QAAGASAVLAKAKAKAEIRILAAALTOHNGDAASLTVAEQYSAFSKILAKDSNTIIL 298
QY 299 PSNPGDVTSMAQAMGVYALTKA--PVPGTDPDLSLSSGSSRDVGTGTDASXDEELDVRKMS 356
Db 299 PSNPGDVTSMAQAMGVYALTKA--PVPGTDPDLSLSSGSSRDVGTGTDASXDEELDVRKMS 358

RESULT 8

PCT-US98-12125-272
; Sequence 272, Application PC/TUS9812125
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; NUMBER OF SEQUENCES: 318
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/12125
; FILING DATE: Jan 01, 1990
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PZ008complete
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 272:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 306 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; PCT-US98-12125-272

Query Match 84.5%; Score 1493; DB 1; Length 306;

Best Local Similarity 99.3%; Pred. No. 4.3e-129;
Matches 304; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 51 RMGRFHRIIEPGLNLTIPYLDRIYQSLKEIYINPEOSAVLNDVNTLOIDCVLRLIM 110
Db 1 RMGRFHRIIEPGLNLTIPYLDRIYQSLKEIYINPEOSAVLNDVNTLOIDCVLRLIM 60
QY 111 DPKASYGEDEPYAVTOLAQTMRSELGKSLXKDYFRERESINASIVDAINCAADCGMI 170
Db 61 DPKASYGEDEPYAVTOLAQTMRSELGKSLXKDYFRERESINASIVDAINCAADCGMI 120
QY 171 RCLRYEIKIHVPPRKESQMOVEAEERRKRAIVLESEGTRESAINVAEKKQAQIILASE 230
Db 121 RCLRYEIKIHVPPRKESQMOVEAEERRKRAIVLESEGTRESAINVAEKKQAQIILASE 180
QY 231 AEKAEQINQAAGASAVLAKAKAKAEIRILAAALTOHNGDAASLTVAEQYSAFSKILA 290
Db 181 AEKAEQINQAAGASAVLAKAKAKAEIRILAAALTOHNGDAASLTVAEQYSAFSKILA 240
QY 291 KDSNTILPSNPGDVTSMAQAMGVYALTKA--PVPGTDPDLSLSSGSSRDVGTGTDASXDEEL 350

Db 241 KDSNTILPSNPGDVTSMAQAMGVYALTKA--PVPGTDPDLSLSSGSSRDVGTGTDASXDEEL 300
QY 351 DRYKMS 356
Db 301 DRYKMS 306

RESULT 9

US-09-209-462B-411
; Sequence 411, Application US/09209462B
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 86 Human Secreted Proteins
; FILE REFERENCE: PZ008P1
; CURRENT APPLICATION NUMBER: US/09/209,462B
; CURRENT FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: PCT/US98/12125
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/049,547
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,548
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,549
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,550
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,566
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,606
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,607
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,608
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,609
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,610
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/049,611
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/050,901
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/052,989
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058,665
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,668
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,669
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,750
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,971
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,972
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,975
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/060,834
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,841
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,844
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/060,865
; PRIOR FILING DATE: 1997-10-02
; PRIOR APPLICATION NUMBER: 60/061,059
; PRIOR FILING DATE: 1997-10-02

```
; PRIOR APPLICATION NUMBER: 60/061,060
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 737
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 411
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-209-462B-411

Query Match      84.5%: Score 1493; DB 16; Length 306;
Best Local Similarity 99.3%: Pred. No. 4,3e-129;
Matches 304; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 51 MGRFHRILPEGLNIIIPVLDRIIRYVOSLKEIYINPEQSAVTLDNVTLOIDGVLRLIM 110
DB 1 MGRFHRILPEGLNIIIPVLDRIIRYVOSLKEIYINPEQSAVTLDNVTLOIDGVLRLIM 60
OY 111 DPKASYGVDEPEYAVTQLAQTMRSELGKLSXDKYFRERESLNASTVDAINQADCMGI 170
DB 61 DPKASYGVDEPEYAVTQLAQTMRSELGKLSXDKYFRERESLNASTVDAINQADCMGI 120
OY 171 CLRREIKDHYPPRVKESOMQVEARRKRAIVLESEGTRESAIVNAEGKKOQILASE 230
DB 121 CLRREIKDHYPPRVKESOMQVEARRKRAIVLESEGTRESAIVNAEGKKOQILASE 180
OY 231 AEKAEQINQAGEASAVLAKAKAKAEAIRILAAALTOHNGDAAASLTVAEQYVSASFSLA 290
DB 181 AEKAEQINQAGEASAVLAKAKAKAEAIRILAAALTOHNGDAAASLTVAEQYVSASFSLA 240
OY 291 KDSNTILPSNPGDVTSMVAQAMGVYGLTKAPVPGTSPSLSSGSSRDVGTDAKDDEL 350
DB 241 KDSNTILPSNPGDVTSMVAQAMGVYGLTKAPVPGTSPSLSSGSSRDVGTDAKDDEL 300
OY 351 DRVKMS 356
DB 301 DRVKMS 306

RESULT 10
; Sequence 36, Application US/09298733
; GENERAL INFORMATION:
; APPLICANT: Valenzuela, Dario
; APPLICANT: Yuan, Olive
; APPLICANT: Hoffman, Heidi
; APPLICANT: Hall, Jeff
; APPLICANT: Rapielko, Peter
; APPLICANT: Adams, R. Mark
; APPLICANT: Agostino, Michael J.
; APPLICANT: Fechtel, Kim
; APPLICANT: Howes, Steve H.
; TITLE OF INVENTION: Secreted Proteins and Polynucleotides
; FILE REFERENCE: AG199-02PM
; CURRENT FILING DATE: US/09/298,733
; EARLIER FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: 60/082,961
; EARLIER FILING DATE: 1998-04-24
; EARLIER APPLICATION NUMBER: 60/086,402
; EARLIER FILING DATE: 1998-05-22
; EARLIER APPLICATION NUMBER: 60/088,994
; EARLIER FILING DATE: 1998-06-11
; EARLIER APPLICATION NUMBER: 60/089,163
; EARLIER FILING DATE: 1998-06-12
; EARLIER APPLICATION NUMBER: 60/091,619
; EARLIER FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36
; LENGTH: 305
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-09-298-733-36

Query Match      84.2%: Score 1488; DB 16; Length 305;
Best Local Similarity 99.3%: Pred. No. 1,2e-128;
Matches 303; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 52 MGRFHRILPEGLNIIIPVLDRIIRYVOSLKEIYINPEQSAVTLDNVTLOIDGVLRLIM 111
DB 1 MGRFHRILPEGLNIIIPVLDRIIRYVOSLKEIYINPEQSAVTLDNVTLOIDGVLRLIM 60
OY 112 PKRASYGVDEPEYAVTQLAQTMRSELGKLSXDKYFRERESLNASTVDAINQADCMGI 171
DB 61 PKRASYGVDEPEYAVTQLAQTMRSELGKLSXDKYFRERESLNASTVDAINQADCMGI 120
OY 172 CLRREIKDHYPPRVKESOMQVEARRKRAIVLESEGTRESAIVNAEGKKOQILASE 231
DB 121 CLRREIKDHYPPRVKESOMQVEARRKRAIVLESEGTRESAIVNAEGKKOQILASE 180
OY 232 EKAQEQINQAGEASAVLAKAKAKAEAIRILAAALTOHNGDAAASLTVAEQYVSASFSLA 291
DB 181 EKAQEQINQAGEASAVLAKAKAKAEAIRILAAALTOHNGDAAASLTVAEQYVSASFSLA 240
OY 292 DSNFTILPSNPGDVTSMVAQAMGVYGLTKAPVPGTSPSLSSGSSRDVGTDAKDDEL 351
DB 241 DSNFTILPSNPGDVTSMVAQAMGVYGLTKAPVPGTSPSLSSGSSRDVGTDAKDDEL 300
OY 352 RVKMS 356
DB 301 RVKMS 305

RESULT 11
; Sequence 36, Application US/09298733A
; GENERAL INFORMATION:
; APPLICANT: Valenzuela, Dario
; APPLICANT: Yuan, Olive
; APPLICANT: Hoffmann, Heidi
; APPLICANT: Hall, Jeff
; APPLICANT: Rapielko, Peter
; TITLE OF INVENTION: Secreted Proteins and Polynucleotides
; FILE REFERENCE: 1966.1009-000
; CURRENT FILING DATE: US/09/298,733A
; CURRENT FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: 60/082,961
; PRIOR FILING DATE: 1998-04-24
; PRIOR APPLICATION NUMBER: 60/086,402
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/088,994
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/089,163
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/091,619
; PRIOR FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-298-733A-36

Query Match      84.2%: Score 1488; DB 16; Length 305;
Best Local Similarity 99.3%: Pred. No. 1,2e-128;
Matches 303; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 52 MGRFHRILPEGLNIIIPVLDRIIRYVOSLKEIYINPEQSAVTLDNVTLOIDGVLRLIM 111
DB 1 MGRFHRILPEGLNIIIPVLDRIIRYVOSLKEIYINPEQSAVTLDNVTLOIDGVLRLIM 60
```


Qy	112	PKKASYGEDEBEVAVTOLAQOTMTSELGXLSXDKVFFERRESLNASIYDAIINQAAADCIGIR	171
Dp	61	PKKASYGEDEBEVAVTOLAQOTMTSELGXLSLDPKVFERRESLNASIVDAIINQAAADCIGIR	120
Qy	172	CLFREIKDINHPPRKESMOMQOVAERBRKRATVLESGETRESAINVAEGRKQAOILASEA	231
Dp	121	CLFREIKDINHPPRKESMOMQOVAERBRKRATVLESGETRESAINVAEGRKQAOILASEA	180
Qy	232	EKAQDINOAGEASAVLAKAKAKAEARILLAALTOHNGDAASLYVAEOYSNFSFKLAK	291
Dp	181	EKAQDINOAGEASAVLAKAKAKAEARILLAALTOHNGDAASLYVAEOYSNFSFKLAK	240
Qy	292	DSNTILLPNSNGDVTSMVAQAMGVYGAULTKAPVPGTDPDSLSSGSSRPDQGTIDASXDEELD	351
Dp	241	DSNTILLPNSNGDVTSMVAQAMGVYGAULTKAPVPGTDPDSLSSGSSRPDQGTIDASXDEELD	300
Qy	352	RYKKS 356	
Dp	301	RYKKS 305	

```

RESULT 12
US-09-723-594-36
; Sequence 36, Application US/09723594
; GENERAL INFORMATION:
; APPLICANT: Valenzuela, Dario B.
; APPLICANT: Yuan, Olive
; APPLICANT: Hoffmann, Heidi
; APPLICANT: Hall, Jeff
; APPLICANT: Rapleto, Peter
; TITLE OF INVENTION: Secreted Proteins and Polynucleotides
; TITLE OF INVENTION: Encoding Them
; FILE REFERENCE: 1966.1009-002
; CURRENT APPLICATION NUMBER: US/09/723,594
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/298,733
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: 60/082,961
; PRIOR FILING DATE: 1998-04-24
; PRIOR APPLICATION NUMBER: 60/086,402
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/088,994
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/089,163
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/091,619
; PRIOR FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-723-594-36

```

Query Match	84.2%	Score 1488;	DB 21;	Length 305;
Best Local Similarity	99.3%	Pred. No. 1.2e-128;		
Matches 303; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

Oy	52	MGRFRILIEPCLNTLIPVLBDEIRYVOSSKEIVINVPQSNATLDMNLOIDGUYLXIRMD	111
Db	1	MGRFRILIEPCLNTLIPVLBDEIRYVOSSKEIVINVPQSNATLDMNLOIDGUYLXIRMD	60
Oy	112	PKKSYGVEDEBYAVTOLAQOTMMSSELGKLSXDKVFRFRBSLNASIYDAINQAADCGIR	171
Db	61	PKKSYGVEDEBYAVTOLAQOTMMSSELGKLSXDKVFRFRBSLNASIYDAINQAADCGIR	120
Oy	172	CLUREIKOIHVPVKRESMOMQVEAERRKATVLESEGTRESAINVAEGKKCAQILNSEA	231
Db	121	CLUREIKOIHVPVKRESMOMQVEAERRKATVLESEGTRESAINVAEGKKCAQILNSEA	180

QY	232	EKAEDINQAGASAVLAKAKAKAEATITLLAALTOHNGDAASLTVAEQVSAFSKLAK	231
Db	181	EKAEDINQAGASAVLAKAKAKAEATITLLAALTOHNGDMAASLTVAEQVSAFSKLAK	240
QY	292	DSNTILLSPEDVTSMTAAQAMGYGALTAKPVPVGTSPSSSGSSRVOGCTDASXDEELD	351
Db	241	DSNTILLSPEDVTSMTAAQAMGYGALTAKPVPVGTSPSSSGSSRVOGCTDASLDEELD	300
QY	352	RYKMS	356
Db	301	RYKMS	305

```
US-09-724-497-36      RESULT    1.3
; Sequence 36, Application US/09724497
; GENERAL INFORMATION:
; APPLICANT: Valenzuela, Dario B.
; APPLICANT: Yuan, Olive
; APPLICANT: Hoffmann, Heidi
; APPLICANT: Hall, Jeff
; APPLICANT: Rapiejko, Peter
; TITLE OF INVENTION: Secreted Proteins and Polynucleotides
; FILE REFERENCE: Encoding Them
; CURRENT APPLICATION NUMBER: US/09/724,497
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/298,733
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: 60/082,961
; PRIOR FILING DATE: 1998-04-24
; PRIOR APPLICATION NUMBER: 60/086,402
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/088,994
; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/089,163
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/091,619
; PRIOR FILING DATE: 1998-07-02
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 36
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-497-36
```

Query Match	84.2%	Score 1488;	DB 21;	Length 305;
Best Local Similarity	99.3%	Pred. No. 1.2e-128;		
Matches 303; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

Qy	52	MORFRRIIEPGNIIILPDLRIARYQSJKEIYINVPBDSAYTLDNWTLQIDGVYLRIMD	111
Db	1	MORFRRIIEPGNIIILPDLRIARYQSJKEIYINVPBDSAYTLDNWTLQIDGVYLRIMD	60
Qy	112	PKASYGVEDPEYAVTOLAQTMRSELGKLSXDVFERESELSNAYDAINQAADCWGIR	171
Db	61	PKASYGVEDPEYAVTOLAQTMRSELGKLSXDVFERESELSNAYDAINQAADCWGIR	120
Qy	172	CARYEIKDIIHVPPRKESMOMOVEERKRKRYTVEESETRESAINVAEGKKOAOIILSEA	231
Db	121	CARYEIKDIIHVPPRKESMOMOVEERKRKRYTVEESETRESAINVAEGKKOAOIILSEA	180
Qy	232	EKAEOINOAGEBASAVLAKAKAKAEAIRILLAALTLQNHGDDAAASLJVAEOYVSASFSLAK	291
Db	181	EKAEOINOAGEBASAVLAKAKAKAEAIRILLAALTLQNHGDDAAASLJVAEOYVSASFSLAK	240
Qy	292	DSNTLLTSLNPEDEVYSMAOAMGYGALTTPKAPVGETPSTLSSGSSROYGOTDASXDELD	351
Db	241	DSNTLLTSLNPEDEVYSMAOAMGYGALTTPKAPVGETPSTLSSGSSROYGOTDASXDELD	300
Qy	352	RYKMS	356

Db 301 RVKMS 305

RESULT 14
US-09-614-150-5415

```

; GENERAL INFORMATION:
; SEQUENCE 5415, Application US/09614150
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CLO00728
; CURRENT APPLICATION NUMBER: US/09/614,150
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5415
; LENGTH: 323
; TYPE: PRT
; ORGANISM: DROSOPHILA
; US-09-614-150-5415

```

Query Match 60.2%; Score 1064.5; DB 20; Length 323;
Best Local Similarity 63.7%; Pred. No. 2,1e-89;
Matches 209; Conservative 55; Mismatches 47; Indels 17; Gaps 2;

```

QY 39 LFVPOQEAAMVVERMGFRHRLLEPGNIIIPVLRIRYQSLKEIYINPEQSAVTLDNVT 98
   |||||
Db 1 MFVPOQEAAMVVERMGFRHRLIDPGNIIIPVADKIKYQSLKEIADVPKOSATISDNVT 60
QY 99 LQIDGVLYLRIMDPYKASYGVDEPEYAVTQLAQTTMRSELGKLSXDKYFRERESLNASTV 158
   |||||
Db 61 LSIDGVLYLRIMDPYKASYGVDEPEYAVTQLAQTTMRSELGKMSMDKYFRERESLNASTV 120
QY 159 DAINGAOCWGRCLRYEIKDHYPRKESQMOVEAEERRKRAIVLESEGTRESAIVNA 218
   |||||
Db 121 DSINKASAMGJACLRYEIRIDRLPTRVHEAMQOMOVEAEERRKRAILLESEGVREAEINIA 180
QY 219 EGKQAOILASEAEKAEQINOAGASAVLAKAKAEAIRILAALTOHNGDAASLTIV 278
   |||||
Db 181 EGKRSRILASEAEKAEQINOAGASAVLAKAKAEAIRILAALTOHNGDAASLTIV 240
QY 279 AEQYVSAFSLAKDSNTLLPSNPGDVTSMVAQAMGVYGALTKAPVPGTSDSLSSGSSRD 338
   |||||
Db 241 AEQYIGAFKRLAKTNTNMTILPSNPGDVNGFVAQALAVYNHVSNS-----NQATKSSSEN 293
QY 339 VQGTDA-----SXDELDRYKMS 356
   |||||
Db 294 VKGVGACLNANSVEYKEIQEDKSSVKMN 321

```

RESULT 15
US-60-191-637-5433
; Sequence 5433, Application US/60191637

```

; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
; TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CLO00392
; CURRENT APPLICATION NUMBER: US/60/191,637
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 42660
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5433
; LENGTH: 323
; TYPE: PRT
; ORGANISM: DROSOPHILA
; US-60-191-637-5433

```

Query Match 60.2%; Score 1064.5; DB 26; Length 323;
Best Local Similarity 63.7%; Pred. No. 2,1e-89;
Matches 209; Conservative 55; Mismatches 47; Indels 17; Gaps 2;

```

QY 39 LFVPOQEAAMVVERMGFRHRLLEPGNIIIPVLRIRYQSLKEIYINPEQSAVTLDNVT 98
   |||||
Db 1 MFVPOQEAAMVVERMGFRHRLIDPGNIIIPVADKIKYQSLKEIADVPKOSATISDNVT 60
QY 99 LQIDGVLYLRIMDPYKASYGVDEPEYAVTQLAQTTMRSELGKLSXDKYFRERESLNASTV 158
   |||||
Db 61 LSIDGVLYLRIMDPYKASYGVDEPEYAVTQLAQTTMRSELGKMSMDKYFRERESLNASTV 120
QY 159 DAINGAOCWGRCLRYEIKDHYPRKESQMOVEAEERRKRAIVLESEGTRESAIVNA 218
   |||||
Db 121 DSINKASAMGJACLRYEIRIDRLPTRVHEAMQOMOVEAEERRKRAILLESEGVREAEINIA 180
QY 219 EGKQAOILASEAEKAEQINOAGASAVLAKAKAEAIRILAALTOHNGDAASLTIV 278
   |||||
Db 181 EGKRSRILASEAEKAEQINOAGASAVLAKAKAEAIRILAALTOHNGDAASLTIV 240
QY 279 AEQYVSAFSLAKDSNTLLPSNPGDVTSMVAQAMGVYGALTKAPVPGTSDSLSSGSSRD 338
   |||||
Db 241 AEQYIGAFKRLAKTNTNMTILPSNPGDVNGFVAQALAVYNHVSNS-----NQATKSSSEN 293
QY 339 VQGTDA-----SXDELDRYKMS 356
   |||||
Db 294 VKGVGACLNANSVEYKEIQEDKSSVKMN 321

```

Search completed: September 22, 2002, 18:26:25
Job time: 728 sec

